Buying Research

A Customers´ Guide

Jens Aagaard-Hansen

P. Stanley Yoder

10378
10878

Community Health Cell
Library and Information Centre
359, "Srinivasa Nilaya"
Jakkasandra 1st Main,
1st Block, Koramangala,
BANGALORE - 560 034.
Ph : 2553 15 18 / 2552 5372
e-mail : chc@sochara.org

COMMUNITY HEALTH CELL
LIBRARY
AND
INFORMATION
CENTRE
BANGALORE 34

BUYING RESEARCH
A CUSTOMERS' GUIDE

COMMUNITY HEALTH CELL
Library and Information Centre
No. 367, Srinivasa Nilaya, Jakkasandra,
I Main, I Block, Koramangala, Bangalore - 560 034.

THIS BOOK MUST BE RETURNED BY THE DATE LAST STAMPED		

Printed by: DBL – Centre for Health Research and Development, Department of Veterinary Pathobiology, Faculty of Life Sciences, University of Copenhagen. 2007

Aagaard-Hansen & Yoder © copyright

ISBN: 87-91521-06-8

RS-100

TABLE OF CONTENTS

ACKNOWLEDGEMENTS v

CHAPTER 1 1
Introduction

CHAPTER 2 5
Asking the right question

CHAPTER 3 15
Choosing the methods for collecting information

CHAPTER 4 39
Analyzing the data

CHAPTER 5 51
Reaching conclusions from the data

CHAPTER 6 57
Presenting the conclusions in the appropriate form

CHAPTER 7 61
Arriving at the right price

CHAPTER 8 65
General considerations

ANNEX A Checklist for rapid check of research proposals 71

ANNEX B List of definitions 73

ANNEX C Further reading 77

ANNEX D References 81

ANNEX E List of abbreviations 83

ANNEX F Format for research proposal 85

ACKNOWLEDGEMENTS

This book builds on our experiences as researchers and teachers. We are thankful to our colleagues, study participants and students for teaching us over the years. During the writing process we have benefited from colleagues thoughtful comments. We are grateful to Drs Sian Clarke, Maria Vang Johansen, Annette Olsen and Birgitte Jyding Vennervald. Thanks also to Helle Lohmann Schøler for technical editing. This book has been under way for a very long time. Almost ten years went from the first idea occurred in Chicago till the publication is now in print. We hope that it has been worth waiting for.

CHAPTER ONE

INTRODUCTION

This document was written to help people get a better understanding of research. The readers we have in mind are program directors and policy makers working in health and development projects. Such people are often asked to organize or approve research on indicators telling how their program has been doing. Or the task may be to carry out a needs assessment: finding out which health problem to deal with or which group to assist.

Without experience in doing research, how can you distinguish good research questions from poor ones, strong designs from weak ones? This book will help you do just that. The examples are drawn from the health sector, but most of the principles apply to other sectors as well. Thus, we trust that professionals from agriculture, education and many other sectors can benefit from reading the book. Furthermore, we deal mostly with research in developing countries. Nevertheless, the principles apply to research in any society.

Program directors often do not have the time, training or experience to conduct research themselves, so they must rely on colleagues and consultants to do the work and interpret the results. In order to effectively choose and supervise such experts, it would be helpful if they understood certain broad principles of the research process and have realistic expectations of what it can do for them. A clearer understanding of the research process will help program managers select the right consultants, ask them better questions, and give more direct instructions about the results they seek.

We have structured this book to follow the way the different phases of a research project unfolds. Research is a process of asking a question and finding the answers in a reliable and systematic way. We can ask any sort of question. Some questions can be transformed into research objectives that can guide the research process. Others cannot. In any case, research projects usually begin with a general question someone wants answered. The text presents the various steps people usually follow in moving from a general question on to the discovery of answers and articulating the implications of the results. The process of moving from a general statement concerning what we would like to learn about a particular population to the formulation of specific research questions to guide data collection and analysis requires experience in all aspects of the research process. This is true whether we are seeking information about the meaning of events to a local population, or collecting information that will be analysed with statistics.

The research process is complex and cumulative. It is complex in that it encompasses so many elements that are linked to each other. And it is cumulative in that the decisions about issues, questions, and methods made early on in the process will determine what can follow. Above all, it is challenging.

This book does not contain detailed guidelines for doing research. There are already plenty of "how to" manuals available about how to conduct all kinds of qualitative research, from focus groups discussions to rapid appraisals and community participatory techniques. There are dozens of books on conducting research in epidemiology and statistics. We believe that most program managers don't need to go into such detail. However, if anybody wants to go more in depth, they can consult the list of books that can serve as a starting point (see annex C).

The key to doing good research lies in the clear formulation of research questions and the choice of appropriate methods for data collection and analysis that match those questions. This book focuses largely on this process. Experience tells us that the efforts that are put into clarifying the questions and thorough initial planning pay great dividends at a later stage. By providing examples of concrete situations to accompany the presentation of general principles, the reader will acquire a better sense of what kinds of things to look for, what sorts of results to expect from different research methods, and some guidelines for the interpretation and presentation of findings.

In addition to the discussions about research questions and methods, we discuss the process of making a budget, considering possible ethical issues, and going through a checklist that a program director can use to verify that all the issues have been addressed in a research design. In order to facilitate the reading of this book we have summarized the main points here (identical to Annex A). These points can serve as a check list for easy reference.

Summary

The questions:

- Is it an appropriate research question?
- Are the objectives clear and well focused?
- Are all the key terms clearly defined?
- Do we have the information already (literature review)?

The methodology:

- Is the study population clearly defined and appropriate?
- Is the design clear and easy to follow?
- Do the data collection methods fit with the objectives?
- Are the procedures for data analysis outlined?
- Is the time schedule realistic?

General considerations:

- Is the budget sufficiently detailed and relevant?
- Are the tasks and the available resources in balance?
- Are the critical assumptions considered?
- Are plans formulated for the dissemination and use of the research findings?
- Is the study ethically sound?
- Is the administrative set-up adequate?
- Does the research team have the required expertise?
- Applying the bird's-eye view – is the proposal consistent?

This summary represents the main points of good quality research. If the program manager keeps them in mind when dealing with the commissioned researchers the outcome is likely to be successful. At the same time, the points reflect the disposition of this book, which is simply an elaboration of this list.

But how does one go about choosing a competent consultant? How do we evaluate evidence of their past work? What sort of questions should one ask of prospective consultants? We do not write about how to find a good consultant. However, being conversant with the basic principles of research helps a great deal in making a short list. Having the ability to assess research questions makes it much more likely that you get the kind of study you need. Understanding the "research way of thinking" and knowing the major pitfalls will enable you to ask critical questions and conduct basic quality control. We believe this book will help you in the process. That it will make you more competent and critical when buying research, and consequently, we hope, a happier customer. We believe the book can be useful for individual readers working with single programs, but it can also be productively used in training workshops. More rarely the program directors have the opportunity (time and funds) to plan and implement a small study themselves. Even in that case this book should be able to guide the research process. The book may be used both as a guide to assess proposals and to conduct research.

CHAPTER TWO

ASKING THE RIGHT QUESTION

What we really want to know

The questions we ask about the natural and the social world around us take many forms. We may ask, for example, how caves were formed in the Shenandoah Valley in the USA where numerous caverns can be visited. Answers can be found in geology text books. We may ask how many patients were admitted to the local hospital last year. For that we should check the local records in the Ministry of Health. We may ask how the HIV virus is transmitted in Zambia. Answers might be found in biomedical and social science research. We may ask how the services at the local government clinic compare to the private clinics in the same area. For that we would interview patients and do observations at clinics. Where we look for answers and how we obtain information depends on the question asked.

For some program directors it is important to know about the disease patterns in a given area. Consequently a clinical epidemiological study is needed. In the field of nutrition, biomedical questions often seek to assess the prevalence of 'stunting' (chronic malnutrition) or 'wasting' (acute malnutrition). Some might want to investigate various micronutrient deficiencies such as iron or iodine or vitamin A among young children. Even if we limit our attention to social interactions, we still have a wide range of types of questions to ask. Do we want to know about knowledge or about practice? If we want to pursue knowledge, are we seeking knowledge in general about a subject (circumcision, cooking, marriage) or knowledge about how people get things done (how does this society organize the circumcision of young boys; how do they prepare meals; how do they organize and perform the rituals that constitute marriage?). If we want to know about practice, is it specific events (the circumcision rituals conducted for adolescent males in a particular community) that interest us, or the image of what constitutes common practice in a region? (what mothers often give their children for breakfast).

If we are interested in specific events and how they occur, we may want to explore the social and ecological context of those events. Individuals share common sense understandings about gender roles, about responsibilities within families and in the public arena, about who should be earning money or attending school, and about who has the right to speak for various groups. These understandings all form part of the picture of the way individuals get things done. Important contextual factors can also be found within health care systems. Questions about the number and quality of health staff available in the formal and informal health care system, the availability of

drugs and consumables at health clinics, and local infrastructure can provide important, contextual information for health planners.

What other researchers have done

Part of the process of designing a research project usually includes conducting a literature search for studies that have asked questions similar to ours. University libraries are often used for doing bibliographical searches, and that remains a productive way to find relevant journal articles. In recent years web based literature search has become a key activity. This has developed into an art in its own right and a description is beyond the limits of this book. Yet studies reported in journals constitute only a small part of the research conducted. The gray literature (reports from donor organizations, Non-Governmental Organizations (NGOs), consultants, government agencies) is far larger than the published literature, but often not accessible. However, many organizations and institutions now have web sites that provide summaries or even complete texts of their reports.

Identifying journal articles and relevant reports is necessary but rather time consuming. Furthermore, it takes ingenuity as well as the right contacts. We look for studies as similar as possible to what we propose to do. Usually, the criteria for the search are place and topic. Have any other studies been done among the Tuareg in Niger where we intend to work? Have any researchers made studies on similar topics (reproductive health, diarrhea, infant immunization) any where else in the world? Or with a little luck, has anybody studied the relevant topic among the ethnic group in question? Maybe the necessary data are already available and a new study will be a waste of resources? Let's consider an example from West Africa.

Example 2.1

A study of complementary feeding of infants in the Mopti area of Mali examined the process of mothers introducing their children less than one year old to foods other than breast milk. We wanted to understand how this process occurs: what mothers notice about the child that might affect this process, how she gets the child to eat, what she gives at what age, etc. We wanted to compare the interaction between mother and child among those with well fed and those with malnourished babies. Among the assumptions made were the following: whether a child is well or poorly nourished depends upon what, how often and how much food the child is given; that there was no food shortage in the household; and that mothers get advice about feeding their infant from their own close relatives and friends.

The literature search turned up dozens of articles about nutrition and malnutrition in West African countries, including ones about populations in Mali. Katherine Dettwyler has published a number of articles as well as a book (1994) on malnutrition in southern Mali. We found that she shared our assumptions and some of our research questions. However, her work does not focus on the mother-child interaction as we wished to do. Therefore, we thought we could continue with our data collection plans. We were not able to find the answers from published reports and articles.

However, even if we are able to find a collection of articles and reports that seem interesting, there are a few more steps to take. We find it useful to look for four elements in articles and reports in order to decide if they may be helpful or not:

- The formulation of the research questions and how they are expressed in the objectives.
- The basic assumptions (which may quite often not be explicit).
- The description of how the data were collected.
- The analytic procedures used to make sense of the data.

Why should it be necessary to go into these rather technical issues as long as the findings are provided? There are at least two reasons. First, if we have been successful in finding documents with findings relevant to our concerns, there could still be methodological aspects (or even flaws) which make the findings unusable. Though we will always look first at the findings of a study, it is difficult to evaluate what the findings are worth if these other elements are not given. Second, even if the results are not of direct use there may still be some methodological aspects which can be borrowed to improve the project which is now being planned for our particular purpose. Have they studied aspects of infant weaning among the Khmer that we can include in our study in Mali? Do they point out some assumptions that we have not

considered? Or have they used some innovative data collection methods which we did not think of?

From research questions to objectives

The statement of the objectives for a research project will flow directly from the research questions already developed. In most cases objectives are articulated at two levels: a general (or overall) objective and a series of specific objectives (more detailed goals). The general objective states either what the research seeks to find or understand about a particular population (e.g. how do women in Mopti introduce complementary foods to their children). In this case of understanding weaning practices, we focus upon how mothers currently wean their children: the choices of foods they have, what foods they give, how they persuade children to eat, and the advice young mothers are given. Objectives are formulated to focus on these issues.

Example 2.2

> Let's go back to the example (2.1) from Mali. The research question focused on the process of introducing solid foods to infants in the Mopti district of Mali. When formulating a derived general objective it could be phrased as "To describe how solid foods are introduced to infants in the Mopti district in Mali". This again may be split into for instance five specific objectives: 1) to describe the kinds of food items used when weaning infants; 2) to describe how these food items are prepared and administered; 3) to describe which persons are involved in the weaning process; 4) to describe the context of the process and the interactions between the persons involved; 5) to explore from where the caretakers responsible for the weaning get their knowledge and skills.

As a rule of thumb the general objective should be equivalent to the 'sum of the specific objectives'. There should not be any major aspects of the general that is not captured by a specific. And there should not be specific objectives that are not encompassed by the general. We recommend that a substantial amount of time is put into the development of the objectives. It is crucial that they be clear and well focused. A study that will investigate "malaria in Zimbabwe" is extremely vague and may well lead to unfocused research. Though it may seem tedious, the wording should be considered carefully. Quite different studies will result if the objective is "to describe the most important diseases among different castes in Tamil Nadu, India" (simply showing what is there), as opposed to the objective of "to demonstrate the influence of caste on the prevalence of diseases" (which implies a causal relationship). A set of clear objectives is the precondition for the rest of the planning. If that task is achieved, the rest falls logically into place. If the objectives are not clear (ambivalent or unfocused) the rest of the planning will be messy as well.

Usually, the research questions reflect the main priorities of the agency that commissions the study. It should be so – that is why they want to invest in a study in the first place. For instance, in a nutrition program that promotes the adoption of a new high protein weaning food, they may wonder why so few women have adopted a new weaning food that the agency has made available.

Asking 'what', 'why' and 'how'

Some public health studies focus on '*what* questions'. Within social science they often want to find out about peoples' knowledge. Do the locals know the ('true') cause of malaria? Do they know the right dosage of a particular drug? Do the traditional birth attendants (TBAs) know the danger signs that should make them transfer the pregnant women to the nearest health facility? Within biomedical research, such questions are very important. What is the incidence of lower respiratory infections among children less than five years of age in Néma, Mauritania? What is the coverage of measles immunization among infants in Tororo District of Uganda? Is there a better effect of the new approach to health education among secondary school children in Kilifi District of Kenya? These questions can all be very important and relevant. But they are also quite narrow. A common example is to find out about peoples' 'knowledge' of a particular area in order to better tailor a health education campaign.

Example 2.3

> Health education forms an important element of public health interventions. Suppose that an NGO is giving high priority to malaria and that prevention rather than curative services are favored. From a biomedical perspective mosquitoes play a central role transmitting the malaria parasites from one person to another. This NGO could be interested in knowing what the target population knows about the role of mosquitoes. Are the mosquitoes connected to malaria? Are fever attacks associated with rain? Or to sorcery? Insights about people's perceptions of mosquitoes can serve as a starting point for designing a health education campaign. Such a campaign would typically aim at encouraging people to eliminate mosquito breeding sites and use bed nets. Giving them knowledge about the role of the insects is assumed to lead to changed behavior.

In cases where we plan to do research in populations that appear quite different from us – where people speak a rare, local language, worship in a religion we know little about, or live and work through social relations rather unlike our own - the questions we ask tend to take a particular turn. When we ask *why* people do certain things and not others, we expect the answers to be found in the fact that they live differently from us in terms of their beliefs, values, religion, and social relations. Biomedical research can also explore *why* questions. Why are nomads living in the Ogaden Desert less prone to suffer form worm infections? Why do women have more side effects of

a given drug than men? Why is the immune response to a given disease related to the season of the year?

Example 2.4

> We were once involved in a study trying to understand why mothers give chloroquine tablets to children with a fever. Why do some mothers give chloroquine tablets to the feverish children while others do not? Among the factors we would expect to be important are: the availability of the tablets, their cost, the history of their use, and the experience of mothers who have given such tablets. These factors could be investigated by interviewing mothers who have given tablets and others who have not, asking about availability, the cost, what people say about these tablets, and their experience in giving tablets to their children. We would also have similar conversations with women who have not recently given such tablets. Our task would then consist of trying to explain the difference in the two groups of women (those who gave tablets, those who did not) from the transcripts of the interviews or conversations conducted.

This sort of reasoning is common in primary health care programs. It goes like this: 1) we identify a problem area, a common condition that we believe could be improved (high rate of malaria); 2) we look at individual actions that contribute to this situation (inappropriate treatment seeking practices); 3) we begin a campaign to inform people about what they should do to treat infants with fever (radio spots); 4) we expect people to change their actions in accordance with the instructions given (seek early treatment).

The drawback to this approach is that it focuses our research too narrowly upon elements related only to what we consider as relevant. We do not examine how parents protect their children from fevers or how mothers respond to fevers. Thus, we will not understand how parents think and act in reaction to fevers in their children. Asking women directly why they did something or not does not usually lead to useful (or at least sufficient) information. It is far better to ask a series of related questions that might reveal how individuals consider an action and what counts for them, and from that point suggest reasons for their actions. Actual reasons for actions (why do you smoke?) always involve a number of factors that change in importance over time. When a person is invited to give only one reason, that reason may or may not be the highest priority at that time. Another and broader way of thinking about questions is asking *how* an event occurs. Let's consider an example related to what happens when a young child has a fever.

Example 2.5

A study could ask the question *how* mothers of young children respond to fever. In this case we would begin by examining the different types of fevers that mothers recognize and the symptoms they notice in deciding what kinds of fever their child has. In many societies in Africa, mothers recognize three or four types of fevers, with only two or three that they consider related to malaria. For instance, among the Toucouleur of Mauritania who live along the Senegal River, mothers recognize two types of fever that are associated with malaria. They distinguish between the two by the severity of the fever and the season in which they occur. Once the local process of diagnosis is clear, we could then examine cases of fever and learn how mothers acted when their infants had fever recently. Often this includes giving a cool bath, giving aspirin or something similar, and perhaps also chloroquine. If the fever is accompanied by convulsions, they may consult a specialist in local medicine for treatment. Finally there may be important non-medical factors (e.g. lack of money or kinship obligations to their mothers-in-law) that influence the way the mothers act. Formulating the questions in this broader manner changes the kinds of information we collect.

The question in example 2.4 related to why mothers use chloroquine tablets can be answered by using appropriate data collection techniques, but the answers are more narrowly defined than in example 2.5. The first set of questions is more typical of public health research, while the second set is more commonly asked by anthropologists. Both of them seem useful ones to answer, but they differ in their assumptions about what information is relevant.

Example 2.6

Another example of the *what* and *why* versus *how* approach comes from research on the use of oral rehydration therapy (ORT) to prevent dehydration of children with diarrhea. The majority of studies of the use of ORT examine the knowledge of individuals about the purpose of rehydration and the preparation of re-hydration fluids. These studies are mostly interested in biomedical criteria for relevance and have not examined how mothers usually deal with cases of diarrhea. Anthropologists have sought to include local understandings of different types of diarrhea and the options for treatment as also relevant (Yoder 1995).

This process of moving from a *what* and a *why* question to asking *how* things get done is often the most effective way to approach these problem areas, for it examines local perceptions as well as those from the outside. Social science studies have a particular ability to answer the broad *how* questions and if they are based on too narrow questions, risk not serving the purpose. However, these different question approaches are most relevant to studies using qualitative methods. In contrast, quantitative (biomedical) studies are mostly dealing with '*what* and *why* questions'. What are the most serious health problems in Mombasa? What are the main causes of

the many cases of maternal death in Columbia? Quite often these questions are asked even more narrowly. How many teenagers are regularly using drugs in Harare? Which proportion of the TB cases in Cape Town is multi-resistant? Though narrow, the questions can be extremely important in guiding policies and interventions. Thus, narrow questions are not bad as such. It depends very much on what purpose they should serve.

Uncovering assumptions and formulating hypotheses

Deciding what we really want to know depends on what we assume to be the 'real truth'. Though we may not always be aware, we all have preconceived ideas about the things we study. Managers have, field workers have and researchers have. It is important to uncover these assumptions as part of the planning phase. The research plan should state what the researchers assume to be true about the situation being studied. The issue of assumptions is closely linked to the various kinds of questions above. The more the study opens up by applying *how* questions instead of *what* and *why* the less vulnerable the study is to preconceived ideas and unwarranted assumptions. However, for biomedical research exploring particular causal relationships, it is necessary to formulate *what* questions.

Example 2.7

Though not derived from research, the following example shows the impact of assumptions. In the 1970s numerous projects trained village health workers (VHW), volunteers chosen by communities, to do health education and distribute simple medications such as first aid, and aspirin and anti-malaria pills. Public health specialists assumed that with proper training and supervision, and a system to supply them with medications, VHWs would spend part of their time each week in responding to basic illnesses in their communities. We assumed that in the spirit of solidarity and cheap labor, these volunteers would continue their work for the long haul. Communities were usually asked to compensate VHWs in some way, but generally not with cash. In most cases this assumption has proved to be false. Not only has it proved extremely difficult to maintain systems of supervision and the resupply of medicines, but VHWs eventually ask to be paid. Our assumptions about the feasibility of this type of intervention were false.

Let's look at another example.

Example 2.8

> Assumptions about the relation between knowledge and practice have also been problematic in primary health care projects. At the beginning of the AIDS epidemic in African countries, biomedical experts made two basic assumptions. First, that most people did not know about the disease and thus people needed to be informed. Second, that once people understood how AIDS was transmitted, and how dangerous it was, people would change their behavior to avoid the virus. The first assumption has proved to be true, for education campaigns have succeeded in informing millions around the world about the existence of AIDS and how it is transmitted. But the second assumption has not been borne out by studies of peoples' response to the risk of AIDS.

In example 2.3 above there was a similar case. An NGO developed a health education campaign, for it wanted to prevent malaria. NGO personnel assumed that knowledge of the crucial role of mosquitoes in the transmission of malaria would motivate people to eliminate breeding sites and use bed nets. It assumed a close relationship between knowledge and practice: as knowledge changes, so will practice. But that assumption has not been supported by evidence from program evaluations. Knowledge does not necessarily lead to changed behavior. The assumption is false (Cancian 1975; Deutscher et al. 1993).

Biomedical studies also make assumptions about disease, the human body and behavior.

Example 2.9

> Suppose that a study is implemented in order to pave the way for a school health project in Tanzania with the particular aim to reduce the burden of some of the important worm diseases (hookworms, roundworms and bilharzia - a popular name for schistosomiasis) among the primary school children. The planners have commissioned a school-based study to assess the prevalence of the worm infections. The research team visits a school in the area and examines pupils in the two upper classes because they are most cooperative. The team assumes that prevalence is the same in all age groups which is not the case. Furthermore, the team assumes that school children are present in school (which is not the case as many are absent due to illness (worm infections included) and other reasons). Finally the researchers don't consider that not all children are enrolled in schools. Maybe 20-30 % of all children are not enrolled? (maybe due to poverty which is associated to worm infections). Hence the bottom line is that because of false assumptions, the prevalence figures will give a skewed picture and a school health program will not reach the most needy.

In designing a study, both the program personnel and the research team already have ideas about what the answers of a given study may be. It is important to articulate those expectations. The process of discussing what we think we will find provides a chance for the planning group to reflect on what they know about the subject and these discussion will help to formulate appropriate objectives. The process of moving from an initial question to more precise objectives amenable to research, gives us the opportunity to examine what we assume to be true about the situation.

Hypotheses are related to assumptions but they are not identical. Whereas assumptions are general and often implicit, hypotheses are particular and explicit. A hypothesis is more like a technical formulation of the likely answer to a question on a causal relationship. If the objective of the research is to compare two interventions, a hypothesis should be formulated accordingly. For instance, it can be hypothesized that "smoking leads to lung cancer", "education reduces the number of traffic accidents" or "drug A provides better cure for pneumonia than drug B" and studies will be designed to address these particular issues.

CHAPTER THREE

CHOOSING THE METHODS FOR COLLECTING INFORMATION

In the previous chapter, we have dealt with the formulation of the basic research question and the way in which it was transformed into one or more research objectives. In order to find the appropriate answer(s), relevant information need to be collected. As already noted, it is to a large extent the research objectives that determine the methodology. There are many different data collection methods, and the trick is to choose those which get closest to the things we want to know, while being efficient and feasible as well.

Quantitative or qualitative?

It would not serve any purpose to formulate yet another set of definitions for these two seemingly contrastive terms. They are often used as adjectives for approaches, analyses, questions, data, method, and research. In most cases the terms refer to the procedures used in looking at information, and the kind of logic used in reaching a conclusion. Research that is called qualitative pays close attention to the meaning of local terms and concepts, while quantitative research may not. Qualitative research asserts that data must be considered in its social context in order to reach a full and valid understanding. Qualitative research produces descriptions of how people get things done. It pays close attention to interaction that happens locally, without outside interference from the researchers.

Quantitative research in health related social science, on the other hand, most often collects information that is expressed as figures and testing for statistical significance of the associations within a large data set. These data sets are produced either by asking the same set of closed questions to large numbers of individuals or by some sort of counting or measuring. Both the questions and the variables rely more on the analytical concepts the researchers use in the research design than on local concepts. Standardizing questionnaires and variables makes it easy to compare data sets, or populations. But to be able to do these calculations, we must first have answers, results that can be combined and counted. Hence the term quantitative is used.

Quite often researchers get into heated arguments about whether quantitative or qualitative methods should be emphasized. So let's deal with that first. We strongly believe that one is not better than the other per se; rather, it is a question of deciding which method will provide data which will reflect most closely the reality we are examining and the use we are going to make of it. Let us give two examples:

Example 3.1

Suppose that the research question that we find interesting is "to examine the height of young men in Chicago". Intuitively, most people would choose a quantitative method to get the appropriate answer to this question. Doing certain measurements would easily give us the desired information: We line the selected group of young men up and start measuring: Robert: 178.5 cm; Jim: 196.2 cm; Peter 210.0 cm, etc. All is well and we can analyze the numbers, take the average, etc. However, in principle, we might have chosen a qualitative method. Suppose that the researcher insists on applying qualitative methods in terms of observation. Then we would get data consisting of statements like: Robert is approximately as tall as my uncle; Jim is about two metres tall; Peter is so tall that he needs to bend his head when his enters the door. All these statements are based on the observer's best judgement, but hardly very useful.

Example 3.2

Another research question could be "to learn about the love life of people between 20 and 30 years of age in Boston". Let us first approach the question in a way that allows us to count answers. One could formulate a question such as: "Are you in love? - Yes, No, Don't know". Or to allow for a bit more variation: "On a scale from one to ten (ten being the highest), how much are you in love?" Other questions could be: "How many times have you been in love?" or "How old were you when you fell in love for the first time?"

Obviously there are many problems with the procedure chosen. Do the respondents have the same perceptions of what it means "to be in love"? How will people feel about being asked such an intimate question by a total stranger? However, in this connection the issue is the extent to which the method for data collection gives a true reflection of the reality. Presumably most people would agree that there is much more to love than what is revealed in the answers to the above questions. Instead let us consider an approach that pays attention to meaning. Let's select some people of the right category and simply interview them in a more open way. A semi-structured interview with questions like: "Tell me how it feels?", Tell me about the first time?" etc. allows for the respondent to chose his or her own words, to tell the whole story for as long as he or she like. They are also free to talk about the issues of importance to themselves.

We have chosen these two examples in order to illustrate that there is no such thing as a good or bad method, but there are methods that are suitable or unsuitable in a given situation. However, it is not an either-or situation. Sometimes (quite often actually), it is an advantage to combine quantitative and qualitative methods. Suppose that we have decided to do research on "unemployment among the inhabitants of Copenhagen". In that case we could learn a lot about the phenomenon from employed and unemployment statistics (e.g. percentages of unemployment in different age groups based on gender, education, residence, etc.). We may also learn a great deal

from qualitative data such as the accounts ('narratives') of some of the people who are unemployed themselves (e.g. their daily routines, their concerns and aspirations, their previous work experience, etc). The right combination of quantitative and qualitative methods improves a study. It gives a more complete picture.

However, permit us some additional comments on the relationship between qualitative and quantitative approaches. First, no research (neither quantitative nor qualitative) gives a absolutely true mirror image of the world. It will always be a partial picture. It will be an approximation, and our role is to make it as complete as possible. Second, the two approaches are more closely related than most people think. Any scientific enterprise, however quantitative, is based on categories, and categories are formulated by human beings as a reflection of current social and cultural trends. The way human beings construct these categories is one of the topics of social science applying qualitative methods. Third, the way we see research is also influenced by socio-cultural perceptions. It is not God-given that statements expressed in figures are necessarily more 'true'. As Keynes once said: "It is better to be approximately right than precisely wrong". So figures can have a false air of reliability (the more decimals the more impressive) in the same way that the many words of social science jargon can hide some very messy concepts indeed.

Are we getting what we are seeking?

Research that relies on individual interviews depends upon the cooperation–good will and good faith–of those interviewed. From time to time interviewees will respond reluctantly or simply lie to an interviewer. If we expect a significant number of people not to speak freely, then we have to consider another method of data collection.

Data collection methods may be inappropriate for reasons other than the respondents' reluctance or outright deceit. Simple lack of memory can be a crucial factor. What did you have for dinner last night? The day before? Five days ago? Most of us don't remember unless dinner was a special event (e.g. a major party). This universal tendency whereby the past is distorted when recalled is called '*recall bias*'. The more ordinary and the more distant an event is, the more it tends to be forgotten. Yesterday's dinner slips into forgetfulness almost immediately, whereas a serious illness episode remains for some time, and the loss of a close relative may never fade.

There is, however, another element to consider regarding the suitability of data collection methods. First an example:

Example 3.3

This example is taken from our work as trainers rather than from a research study. One of us was once teaching a module of "teaching skills" as part of an international course on tropical diseases. The course management had designed the course in such a way that all modules should be evaluated by a set of multiple choice questions (MCQ) given before and after each module in order to see whether the teaching had any impact. We appreciate that MCQs is a good way to quantify specific knowledge. But there is much more to teaching than knowledge. You may know the pedagogic textbooks by heart and still be a lousy teacher. Or vice versa. Consequently, I objected strongly to the MCQ methodology and got the permission of the management to evaluate the students based on their performance in a real life teaching session instead. This evaluation method based on observation got much closer to the skills in question.

Because of its subject, methods of investigation, and ways of reaching conclusions, social science is generally very verbal. A great deal of data is collected through question and answers sequences: in a conversation, through written essays, or from answers given on printed questionnaire forms. However, some social scientists within the main disciplines of psychology, sociology, and anthropology to a large extend rely upon observations, thus producing both visual images as well as text. Social scientists rely so much on texts in part because we study human interaction, and our primary source of data is peoples' accounts of their own or others' actions. Since we do not have the time and other resources to observe many practices, we collect descriptions of them. In addition, we are very interested in hearing about what individuals know and think about nearly any and every subject possible. For that too we need words. We construct verbal images of both events and of relationships. And finally, much of human wisdom can only be passed on to the next generations verbally, which in itself gives a special importance to texts and to narratives.

However, sometimes researchers apply a verbal method though a visual method would clearly get closer to reality. As mentioned in the example 3.3 above, the students were taught skills (how to teach), but were originally tested on knowledge. This is a classical dilemma. Interviewing the study population about what they do instead of verifying by observation has been the method applied in thousand of studies. Don't expect to get completely reliable answers to questions such as: "do you wash your hands every time you go to the toilet?"

Basically, what we are discussing here is the question: are we getting what we are looking for? The technical term is *validity*. Validity is a complex but important concept in the context of research, but it is given so many meanings. The chance of

getting confused is large indeed. In fact, there are two terms that are widely used in research that refer to quality: precision and validity. *Precision* refers to the degree to which any estimate or measurement will be the same over and over. Validity refers to whether or not a measure of something actually measures what we think it does.

Stone and Campbell have published an article comparing the validity of two methods of collecting data about family planning in Nepal.

Example 3.4

> Stone & Campbell (1984) conducted a study in Nepal in order to compare the validity of two methods for collecting data on family planning methods. One method was a duplication of the 'Nepal Fertility Survey', which had been conducted by outsiders from the Ministry of Health based on questionnaires in order to collect data on people's knowledge of various methods of family planning. The other approach was more informal and based on casual conversations and interviews conducted by people known to the locals. The study showed that the latter approach revealed a much higher proportion of knowledge in the study population. The main issue was that topics were very sensitive and not something that people usually discussed with strangers.

The example from Nepal leaves no doubt that the in-depth interviews carried out by social scientists, after having spent time to build rapport with the interviewees, is a better method when it comes to getting personal information about family planning in this rural community. It seems that the interviews simply came closer to the truth. The open-ended nature of the interviews allowed women to draw upon their own concepts and understandings of contraception, rather than responding to questions using concepts new to them. And the alternative methods allowed the women to speak more freely about what to them was a very sensitive topic. The method provided more valid data.

Many other examples could be given. One would presume that a diagnosis made by a well-trained medical doctor on a well-equipped hospital is more valid than that of a quickly trained community health worker in a rural area of Burkina Faso. The age of an individual is assessed in a more valid way by reading a birth certificate (which may also be wrong, by the way) than by counting the wrinkles around the eyes.

Is it operational?

Now and then a given data collection method may be the best in theory, but it cannot be used for one reason or another. For instance, the research question may be very sensitive. The question in example 3.2 regarding "love life" in Boston will presumably be considered rather private in most parts of the world. The most accurate picture possible would be obtained by adding around the clock observations of the research persons for extensive periods of time in addition to the in-depth interviews.

Obviously, this is inconceivable. Except for special television programs, no one will tolerate such observations of their daily lives. However, even the interviews may pose problems. Perhaps some individuals in Boston don't want to talk about these matters to a stranger. In many parts of the world, it would be impossible.

There are many issues which people may find embarrassing to discuss. Not only sexual topics, but also consumption of various substances (e.g. alcohol and tobacco), wealth or level of income (which may be misused by the tax authorities or greedy relatives) may be problematic research topics. It should be borne in mind that sensitive issues vary from one context to another. Whereas most people tend to under-report alcohol consumption, in some groups it is considered highly prestigious to drink which may lead to over-reporting. We may have to accept less reliable, but more feasible indicators such as total sales of alcohol for rates of drinking. Such a measure is called a *proxy*. A proxy is an element that stands for something else in a data collection context. For instance, the Demographic and Health Surveys directed by Macro International commonly use a combination of household objects as a proxy for wealth.

Example 3.5

> The construction of a wealth scale is common for demographic as well as health surveys. For scales such as a wealth scale, the Cronbach alpha score of the scale will likely be cited to indicate how well the units of the scale are inter-correlated. In addition, much can be learned about the assumptions made by the researchers if we examine what they chose as proxies for wealth and how the put the scale together. Most of these scales combine aspects of house construction with ownership of furniture and machines, either appliances or for transportation (bicycle, motorbike, car).

We may also have to discard the best choice of method, because it is either too expensive or simply not operational in a given context. The budget ceiling of a project may not allow the purchase of a piece of equipment (e.g. a video camera or ultrasound equipment). Or the conditions for using a technique are not available (e.g. lack of electricity for a machine to operate). Maybe the appropriately trained manpower is not available (e.g. lack of literate females with sufficient knowledge of a given language). Or the needed time is simply not available.

The main methods for data collection

The methods for data collection vary enormously. Here we have classified them according to the ways the information is obtained. The method appropriate to each situation depends mainly on the way the research questions have been formulated.

Interviews

We think of interviews as composed of two elements: questions and answers. The classification of various interview techniques is very much based on the level of 'control' in these two elements. Is the interviewer given any guidelines about what to ask about? Is he or she even told exactly how to phrase the question? Is the interviewee given a free choice how to answer or is it simply a choice between 'yes' or 'no', or other pre-coded answers? No approach is better or worse per se. It is a question of which interview method is the most appropriate in a given situation. Furthermore, one can interview people individually or in groups. These combinations give the following common interview methods:

Individual

Non-structured, open-ended

This is a dialogue about several selected themes with a minimum of predetermined questions. The interviewee as well as the interviewer can lead the talk in new directions. Points of interest are taken up as the conversation develops. The respondent can answer in any way he or she likes.

Semi-structured, open-ended

This is based on a predetermined question guide that outlines main themes of interest. However, the questions don't need to be asked in a particular way and can be probed when needed. New leads can be followed if desired. The respondent can answer in any way he or she likes.

Questionnaires

Questionnaires are composed of a list of questions designed to be asked in exactly the same way each time, along with answers that are pre-coded. Usually the possible questions are not read out loud to respondents. For a questionnaire to be useful and effective, the questions must be understood in the way intended by the researchers. In addition, researchers must know what most of the responses will be, although they may have no idea about the distribution of the answers. For instance, a question to be asked of African students in Denmark might be: "What is your nationality?" The possible answers could be: 1) Kenyan; 2) Ugandan; 3) Tanzanian; 4) Other. Using predetermined questions and answers for individual interviews effectively requires pre-testing to check on how questions are understood, and to verify that the answers given can be easily marked according to the codes provided. If the answers are pre-coded in a way that does not match the responses that are given, the answers cannot be properly marked. For example, the question cited about "nationality" will not produce useful answers if students come from countries other than east Africa, for all answers would be coded as "Other". Although the code "Other" would be accurate in such a

situation, it would provide very little information, for it would indicate only that they are not from Kenya, Uganda or Tanzania.

Sometimes questionnaires include structured questions that allow open-ended answers, and the interviewer may write the actual answers and code them later. Questionnaires can be administered person to person or they can be given to the respondent who then fills it out him or herself.

Narratives

These are accounts of a sequence of events in someone's life (e.g. a childhood, a career or an illness episode) given by an individual and then set in chronological order by a researcher. Although such accounts are recorded in the words of the person telling the story, the person may not talk about the events in a story in the sequence in which they occurred.

Group

Group interviews

These can be informal group conversations or more formal interviews. The questions can be more or less open and there can be one or more interviewers.

Focus Group Interviews (FGD)

Basically a FGD is a semi-structured, guided group discussion with some additional characteristics. It is focused in the sense that the (8-12) interviewees share some features relevant to the given research question (e.g. women between 20 and 30 years of age who have all given birth at least once), and that certain topics chosen in advance are discussed (hence it is semi-structured). The aim is to facilitate group dynamics in such a way that interaction is taking place directly between the interviewees who may react to each other in various ways. The research team will be represented by a facilitator and a note-taker.

Observations

As with interview techniques, observations can be more or less guided. The observer can be more or less directed to look for specific actions, events or movements. Usually the observed are less involved than in interviews. A wide range of recording techniques are available: video, snapshots, hand-written notes, drawings, filling out check lists. Anthropologists use these techniques along with extensive written notes from observations.

Structured observations

Sometimes it is useful to observe very specific things. How many times does the health worker use the stethoscope when making a diagnosis? Which individuals are in contact with the water source from where they can contract the disease

schistosomiasis? How many people used the public facilities for washing hands when they left the latrines at the sports field? How many marriage parties took place in the village during the month of December? We observe individual movements, social interactions, conversations, and events.

Open observations

As with the open-ended interviews, the researcher has the freedom to direct his or her attention in the direction which seems most appropriate. Unforeseen events (e.g. an accident, a child birth, etc.) may occur and be observed to reveal insights into how people deal with such cases and the roles they play.

Participant observation

Participant observation is a type of observation in which the observer also takes part in the events being observed. One person plays two roles at once: observer of social interaction, and participant as well. Participant observation is seen by many to be the hallmark of anthropology.

Though we have described the methods of observation above in three distinct categories, the categories may not be that distinct in real life. They differ in the degree to which observations are structured or not, and the position of the observer vis-à-vis the observed population (outside, inside, back and forth).

Secondary data

Secondary data is the term for information that has been collected by other people for other purposes than the present study. It is figures, text or visual images which have been produced previously in another context, and often for purposes other than research. Documents useful for research may be statistics from the Health Management Information System (HMIS) of the Ministry of Health, newspaper clippings, school books, health education pamphlets, files of health posts, "road to health cards" of children, to mention just a few. Visual images may be posters used for health information, TV spots advertising for pain killers or old family photographs. Secondary data can be useful in either qualitative or quantitative approaches.

'Innovative' methods

We have mentioned above the data collection methods most frequently used in qualitative studies. However, there are many others available. Just to mention a few, there are:

Draw and write

In this technique, usually used with children, the person is asked to make a drawing of a particular concept, e.g. 'a healthy household'. Later on comments are written about the scenes appearing on the paper to explain the images.

Diaries

A person (key informant) is asked to write a diary. For example, a Traditional Birth Attendant (TBA) could be asked to keep a diary on the clients that came to her for assistance.

Essays

A person (e.g. a school child) is asked to write an essay on a particular topic, e.g. 'the causes and treatment of malaria and how it can be prevented'.

Ranking

A group of people is asked to participate in a group discussion on a certain topic while the researcher is present. They are asked to rank certain items in an orderly sequence based on certain criteria. For example, the households of a village can be ranked based on the criterion of 'wealth' (however that is defined in the local context). Or various types of maize can be ranked based on their drought resistance by a group of farmers. In this way the 'items' are put on a scale of more or less wealth, or more or less drought resistance, all in relation to one another.

Theatre and role plays

One or two researchers can implement this method with a group of informants. The variations are many. A small play which describes a situation of relevance to the project (e.g. domestic violence, or alcoholism) may be enacted and then discussed by the spectators. Or a group of the local population with particular characteristics (e.g. young mothers, or village elders) may be asked to portray a typical situation (e.g. a visit to the ANC services or a meeting with the District Health Committee).

'Biomedical methods'

Strictly speaking 'biomedical methods' is not an accurate term. But we wanted to introduce briefly at least a few of the most frequently used research methods. Medical and epidemiological textbooks are full of tests and examination methods for assessing clinical diagnoses. These methods are beyond the scope of this book. We give just a few examples relevant to public health concerns.

Clinical examinations

A clinician typically makes an assessment such as a diagnosis (e.g. pneumonia) or the identification of a symptom (e.g. an enlarged liver) based on a combination of the patient's history and physical or psychic signs. Within the field of nutrition, simple measurements such as height and weight and upper arm circumference provide important indicators of acute and chronic conditions of malnutrition.

Para-clinical examinations

These include the numerous ways in which blood samples (e.g. a blood slide to look for parasites) or stool samples (e.g. for the identification of worm eggs) are used for specific diagnoses. However, in principle, more high tech tools such as CT scans for brain tumors or ultrasound scans for intra-abdominal cysts can also be used, though these are rarely feasible in research in developing countries. In this category we also find various 'biomarkers' (i.e. methods where a certain sign gives an indication of a diagnosis tough it is not completely reliable). A good example is red urine or dipstick examination of blood in urine as an indicator of urinary schistosomiasis infection. Though not quite as reliable as examining the urine for eggs, it is much cheaper and more convenient for field research. In a way red urine can be seen as a proxy for the disease.

Defining the study population

Who are they?

Once the data collection methods have been chosen, we decide from which *population* the information will be collected. That decision should flow naturally from the research questions and the assumptions formulated. The groups are generally defined by key variables that we believe may be significant in differentiating one group from another. These are factors such as age, gender, residence, ethnicity and occupation. If the research objectives are well formulated the study population will be clearly given, e.g. young men of Chicago (example 3.1). In other cases it is slightly more complex. Think of a study of teenage sexual practices in Brazzaville. We would certainly interview teenagers there and observe how they interact. But should we also consult parents and teachers? How about social service personnel? The bar men at the local discotheques? Thus, the first step is to clearly identify from which group(s) of people (populations) to gather information.

The second step is to decide whether all the members of the group(s) in question should be approached (interviewed or 'measured' depending on the choice of method) or whether only a sample is needed or possible. In some instances this choice seems quite obvious. The total population may be so small that it is feasible and natural to

include them all in the study. If our study population is the senators of the US Senate, why not speak to all 100? But if the study population consists of all primary school children of Thailand, this is obviously not feasible or even necessary. In the latter case a *sample* has to be taken out of the total population.

In principle there are two ways of sampling: randomly or purposively. As the term indicates, *random sampling* utilizes the principle of chance. Using some system that does not depend on human calculations (e.g. a software program or a mathematical table), a certain number of individuals are selected as the sample (e.g. an interview). Random sampling is mostly used in studies that rely upon quantitative methods of analysis and relatively large samples.

In contrast, *purposive sampling* is used more often in qualitative studies with smaller sample sizes. The researcher chooses the sample systematically according to certain principles. The researcher sets up the sampling criteria in each study, most often in a way that achieves a maximum variation in the sample population. (e.g. in terms of age, sex, schooling or whatever is relevant to the study). Assumptions about what elements are relevant have a clear impact on this process.

There is no approach to sampling that is best and that fits all situations. As was the case with qualitative and quantitative approaches, it is largely a question of what is most appropriate for the study in question. However, whether the sampling is done randomly or purposively, it must be done systematically. One should always have a clear picture of what is sampled (is it children or is it the households in which they live?). Technically this is referred to as the *sampling unit*. Furthermore, it should be spelled out from within which frame the 'units' are sampled (e.g. the present Senators of USA or the households of Kanyandiri village in Western Kenya). This is termed the *sampling frame*. The list of studies that turned out to be useless because of 'sampling mess' is longer than the Congo river.

Consider the following:

Example 3.6

Suppose that our research objective is "to describe the major illnesses among infants in a province in western Kenya". If the method chosen combines clinical examination and lab tests, the study population will evidently consist of the infants themselves. But what group of infants do we use to select our sample? Depending on whether one examines a sample of infants from a village, from the 'Out Patient Department' or among those admitted to the local hospital, the illness pattern will vary significantly (e.g. one would expect more cases of common cold in the village as opposed to relatively more cases of the serious cerebral malaria at the hospital). If the method of choice is reported illness, the caretakers are the ones to be interviewed.

There are certain pitfalls that must be avoided for the whole project to succeed. First and foremost, there is the issue of *selection bias*. Let's start with an example.

Example 3.7

Let's recall example 3.1 in which we examined "the height of young men in Chicago". As the objective is formulated, it includes all young men within Chicago. Consequently, we need to define what we mean by 'a young man' (let's say 18 to 25 years) and Chicago (let's say the area within the borders of the incorporated city). When the study findings are published, people will expect them to give a true picture of reality. If it is stated that the average height is 191.5 cm and that it ranges from 149 cm to 208 cm, people will expect these figures to reflect the whole population in question.

In order to be certain about this, the researchers need to be very particular about whom they measure. Suppose the research team carry out all their examinations on Michigan Avenue. They set up their research site and ask the first 200 young men between 18 and 25 to allow them to measure their height. This is acceptable as long as all young men in Chicago have an equal chance of passing by that point. But what about the population in one of the poorer suburbs where the nutritional status is presumably worse than among the richer parts of town? Maybe the population there has another ethnic composition, which genetically makes them likely to be either taller or shorter than the rest of the young men in Chicago? Suppose that the team of 'Chicago Bulls' (a famous basketball team) just passes by on Michigan Avenue when the data collection takes place?

If a study has a selection bias, the study will examine a population different from the one we think. This is important when the final conclusions are drawn (see Chapter 5). We think that we can reach a conclusion about all young men in Chicago, but in reality it is only a fraction (the well nourished, the basket ball players, the wealthy). The ones we have examined do not truly represent the whole population. The term *representivity* is often used in this connection. Sometimes a selection bias is never discovered, though the study may come up with findings that apply only to a sub-population. Sometimes it is discovered during the data analysis. The best strategy is, of course, to consider the issue of sample selection carefully during the planning phase in order to avoid unpleasant surprises. Selection bias is a kind of *systematic error*, i.e. it contributes to skewing the research findings in a particular direction.

The issue of *drop-out,* which is of importance especially to quantitative studies, is related to selection bias. Once we have selected a number of individuals from a certain category, they will not all be available. Perhaps they have moved to another address, maybe they are busy working in the field, or maybe they simply don't want to participate. If the study has chosen to make consecutive interviews with people, we may not find them when we come back for the second encounter. There will always be a certain proportion of people unavailable at a certain time, so we should make rules about how they should be replaced by individuals just like them. This should be considered in the planning phase as well as during analysis. Should new individuals

be recruited instead, and if so, according to which criteria? Or should a reduction of the size of the study sample be accepted? How big a drop out rate can be tolerated? We are dealing here with the risk for selection bias again. Suppose that the individuals dropping out belong to a particular category. Is it the poor who are not available because they are busy working in the fields? Is it people from a particular congregation which forbids them to participate in clinical studies?

How many should we select?

Having clearly defined the study population, the next step is to consider the sample size. How many questionnaires should be administered in order to reach reliable conclusions? In some cases the decision is easy. The total population may be so small that it is feasible and natural to include them all in the study. In most cases it is not feasible or even necessary and a sample has to be taken out of the total population. But how many?

In order to get to grips with this important question, we need to digress a bit into the statistical way of thinking. From a research perspective, there is an element of chance in all aspects of life.

Example 3.8

If one tosses a coin there are two possibilities – heads or tails. In this particular example (and provided that the coin is not manipulated in any way) we know the odds – namely, that there is a fifty/fifty chance of getting either of the two. This is the logical reality, but suppose that we want to prove it. The first study has planned to toss the coin ten times. So from a hypothetical perspective, one would expect 5 heads (50%) and 5 tails (50%). But in a real life experiment we may well find a 60% - 40% distribution or 20% - 80% - all because of chance. In the second study we toss the coin a thousand times. In that experiment we are likely to get a result that is much closer to (but probably not exactly) 50% - 50%. This is simply because the chance element in every single tossing is having a smaller impact on the final result the more times the coin is tossed in total. The outcome of every toss gets neutralized by the outcomes of all the others which bring it closer to the average.

In principle, all types of research are operating in this context of chance when sampling occurs. Technically, it is said that there is an element of *random error* in all results. Suppose we conduct a study of HIV infection rates among 20 years old male soldiers in country C and Country D based on samples of the populations of soldiers in the two countries. We find a higher prevalence of HIV infection in country C as compared to country D. This may be correct or not. Maybe the disease is actually more widespread in country C. Maybe it just so happened that the researchers, by chance, sampled more HIV positives in country C although the proportion is just the same in country D. But the opposite can also be the case. Suppose the study showed

that there was no difference between the two groups. This could be a genuine reflection of the truth. But it could also be a false picture based on the chance event that the researchers did not locate the HIV positive cases that were in reality more abundant in country C.

The phenomenon of random error can never be avoided, but it should always be addressed and studies designed in such a way that it is reduced to reasonable proportions. Example 3.8 showed that one way of reducing the influence of random error is to increase the sample size, i.e. to examine more individuals out of the total study population. There is a snag, however. What seems to be a large sample size may prove to be rather small. If a questionnaire is administered to a thousand people, it seems like a lot. But if a given variable is going to be analyzed according to the respondents' sex, age group, ethnicity, residence and social class, the sample size is reduced by just as many fractions. Consequently, the way in which the data is going to be analyzed should be considered when the sample size is determined.

In some cases there is another factor which influences the sample size needed. The more outstanding the finding is, the smaller the sample size needed. Let go back to the example of comparing HIV in country C and D. If the true difference is big, a smaller sample size is needed to show it; if the true difference is small, a larger sample size is needed to obtain the same level of certainty.

Sample sizes should be calculated based on the estimated differences expected and the degree of certainty desired. There are plenty of different mathematical formulas used for these calculations, but that is beyond the scope of this book. As mentioned, the issue of chance has to do with random error. This should not be confused with systematic error, in which the basic design of the study has skewed the findings.

Random error introduces an element of imprecision in the study findings, for the role of chance can shape the data in all sorts of ways. Consequently, it reduces the certainty of the conclusions. Example 3.8 describes such a situation. Systematic error, on the other hand, pulls the data in one particular direction. The study findings may be much focused, but unfortunately they are wrong (e.g. too high or low as compared to the truth). When discovered, however, one can sometimes make corrections in the conclusions to allow for this error.

Kirk and Miller (1986:19) have a fine example of the use and meaning of both terms with reference to a thermometer, which we have modified a bit:

Example 3.9

> Suppose we use a thermometer to measure the boiling point of water at sea level. A thermometer that shows the same reading of 82 degrees each time it is plunged into boiling water gives a precise measurement. A second thermometer might give readings over a series of measurements that vary quite a bit around the mean value of 100 degrees. The first thermometer is very precise (showing 82 every time), but does not give a valid result (maybe due to some fabrication fault), whereas the second provides valid measurements (around 100 degrees) while being rather imprecise.

There is more to random error than stated here. In chapter 4 we will go more into detail on the concepts of *significance*, *confidence intervals* and *power* which are used in various ways to quantify the risk for random error.

Although the terminology related to random error is closely tied to quantitative research, it is also relevant to qualitative research, though it is usually not dealt with explicitly. Chance can also play a role in the process of interviewing. If the researcher conducts an interview with one individual only, there is a higher risk that the conclusions drawn do not reflect the perceptions of the community at large. If 100 people from the same community are interviewed, the conclusions will rest on a much firmer foundation.

Designing a plan for specific activities

Any study must have a *design*. Apart from describing the data collection methods and sampling procedures, it consists of a plan outlining what should happen and in which sequence of time. Should the same kind of data be collected once or several times? Should all data be collected at one point in time or should it extend over a period of time? Do we plan to follow the study population into the future, or do we include data collected (maybe by somebody else) in the past? Should the researchers introduce an intervention (e.g. a drug) as part of the study? All these issues are important and should appear clearly in the time schedule pictured in the study proposal. This is where the study's feasibility can be assessed in terms of time consumption and the logic of the sequence of events. Obviously, it also gives an opportunity to cross-check with the budget (see Chapter 7).

Example 3.10

Suppose that a number of different qualitative data collection methods chosen to study how mothers of young children respond to fever as outlined in example 2.5. The researchers have suggested using open-ended interviews and focus group discussions (to learn about the different options) and a questionnaire survey (to estimate the frequency of the various practices in a broader population). Here it makes sense to conduct the survey last, since the questionnaire will contain of a number of pre-defined options for answers. The individual interviews will provide information for formulating the possible answers. The interviews and FGDs are prerequisites for formulating the right options and a time schedule reveals whether the researchers are aware of that.

Description or causal relationships?

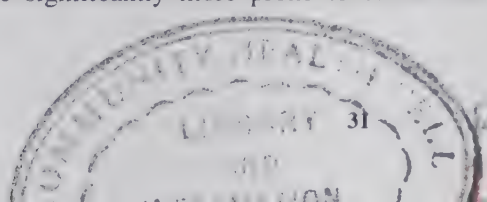
Basically, there are two kinds of studies, those which are mainly descriptive, and those which set out to explore causal relationships.

A *descriptive study* aims at describing what is there. Some of the examples mentioned above are descriptive, e.g. "the examination of the height of young men in Chicago" (example 3.1) and the study on "the love life of people between 20 and 30 years of age in Boston" (example 3.2). As it appears, a descriptive study can be either quantitative or qualitative (or any combination of the two). Usually a study describes more than one 'thing' (*variable* in quantitative jargon).

Example 3.11

Health planners everywhere need evidence on which to base their planning. Basic information on "the major diseases in Zinder province of Niger" is a prerequisite for rational allocation of resources. Should the health care system provide drugs for malaria, surgeons for operations of lung cancer or health information regarding safe water and sanitation? Which parts of the population are suffering from which diseases? The pregnant women? The children belonging to a specific ethnic group? The low income groups? Descriptive data are necessary in order to utilize the scarce funds and manpower in the most cost-effective way. In a well functioning health care system, collection of such information should be a continuous process. As this rarely takes place, such studies are necessary.

Suppose that such a study has been carried out. The health planners now have a data set at their disposal containing a lot of information. Depending on the scope and degree of details of the examinations and questions of the study, it will be possible to show links between various variables. We might find in the example 3.11 above that "pregnant women from the Balese Ndake ethnic group are significantly more prone to suffer from malaria than those from the Keliko group" or that "children from poor families are significantly more prone to suffer from diarrhoea than the wealthy".



Technically, these links are called "*correlations*" referring to the fact that two variables co-vary. Either they both increase or decrease together or one goes up when the other goes down. But (and this is crucial) we can not infer anything about causality.

Cross-sectional studies are quite often referred to as descriptive studies because all the data are collected at one specific point of time, and do not extend either backwards into the past or forwards into the future. More precisely, a cross-sectional study is carried out more or less in one go (although maybe during a long time), and it is not interrupted by related research activities such as interventions, necessary brakes, etc. A cross-sectional study is the simplest design and consequently relatively cheaper and easier to carry out. But its findings are limited in the sense that we cannot draw conclusions about causal relations. If a project wants to study causality, it needs data from two points in time.

For example, based on the finding that "children from poor families are significantly more prone to suffer from diarrhoea" we are not justified to say whether the children get diarrhoea because they live in poor families (who can't afford a latrine and clean food) or whether the families are poor because of the children's diarrhoea (they have to spend a lot of money on treatment for their children). Or maybe there are one or more underlying causes that determine both the childhood diarrhoea and the poverty (e.g. low educational standard of the mothers). If the research question aims at throwing light on some causal relationship, another (and more complex) design should be chosen.

Evidently, a causal relationship signifies that one factor can be seen as the cause of another. From a public health perspective, it can lead to something negative (e.g. lack of clean water as the cause of diarrhoea) in which case the cause is seen as a *risk factor*. Or it can lead to something good, e.g. attending ante-natal-care (ANC) clinics as a means of reducing maternal mortality. These findings are of direct relevance to health care, either at the individual level, where a clinical worker should know what harms or cures a given patient, or at the public health level, where the health planners should know which interventions are available in order to address particular health problems. So how do we design research studies to investigate causal relations?

Example 3.12

The many harmful effects of smoking have now been well documented. But suppose that we wanted once more to show that you can get lung cancer from smoking. We would then design a study where a group of people over a period of time is first exposed to the risk (tobacco) and later on determine how many (what proportion) get lung cancer. We then compare that result with results from non-smokers. The research can be done in two ways with regard to time: retrospectively or prospectively. The researchers can identify lung cancer patients and trace into the past (retrospective) whether they have been exposed to tobacco. Or they can identify smokers today and follow them into the future (prospective) waiting to see how many develop lung cancer. It is, of course, much more complex than that, but we want to highlight here the importance of the time element in the study design.

Let's consider another example:

Example 3.13

The pharmaceutical companies are constantly on the outlook for new drugs to cure various diseases. The technical aspects are extremely complex, from the initial chemical screening of potentially promising substances, via animal experiments, to trials on actual patients. In order to prove that a new drug (A) is better than the old one (B), a study has to be designed to compare the effects of drug A and drug B on two identical groups of patients. One group is given drug A, while the other is given drug B. Following the patients over time looking for the effects of drugs A and B will provide evidence for saying whether one drug is better than the other. This kind of design does not apply to pharmaceuticals only. The same kind of study would be necessary in order to prove the effectiveness of health education, bicycle helmets, new methods of family planning, etc.

The two examples (3.12 and 3.13) raise quite a few issues of importance to study design. It is necessary to be familiar with the following concepts in order to be able to assess study proposals. Let's take one at a time.

All studies dealing with causality contain an element of comparison. Is drug A better than drug B? Is health education on the risk factors of a particular disease better than no health education? Are smokers more prone to get lung cancer than are non-smokers? A relevant concept here is *control group*. The non-smokers serve as a control group. The ones getting the old drug B serve as the control group for the ones getting drug A. However, in order to get a correct answer, the researchers need to make sure the two groups of people are similar in all respects apart from the difference under study. It is crucial that the smoker and the non-smoker groups have the same composition in terms of age, sex, exposure to other relevant substances (e.g. uranium), etc. If this is not ascertained, it can lead to seriously erroneous conclusions (e.g. that drug A is falsely proven to be better than B or that an actual advantage of

drug A is not coming out clearly). In fact it is the old ghost of selection bias coming up again. The two groups are not selected from the same population (are not representative of the same population) and consequently, cannot be compared. But if the researchers are not aware of that, all sorts of false inferences can be made.

In example 3.12 and 3.13, we mentioned that researchers need a time span in to order show the existence of a causal link, and that it can project either into the past or into the future. If we combine the time perspective with the control group issue, we get two basic designs: the case-control study and the cohort study.

The *case-control study* is built on a comparison between a group of 'cases' (typically individuals suffering from a given disease) and a control group (which is supposed to be similar to the other group in all respects except for the disease). Hence the term 'case-control'. Such a study is *retrospective*, i.e. it looks into the past. In principle, it can look for potentially harmful or beneficial effects. In the case of lung cancer (example 3.12), smoking might be that harmful factor, leading us to formulate the hypothesis that "among the cases of lung cancer which have been diagnosed today, a look into the past will show that they have smoked significantly more than those of the control group". However, another hypothesis could be that "the individuals in the control group (not having cancer) had been exposed significantly more to high quality health education". Common to both of these examples of case-control studies is the fact that they are designed in such a way that the absence or presence of a diagnosis today is taken as a starting point. The researchers then either go hunting for data, which have been collected for some other purpose, or base their data collection on the memory of relevant persons.

The other type of study is called a *cohort study*. In such a study, the research population is split into two according to whether they are presently exposed to a given factor or not. In principle it could be anything – potentially harmful or beneficial. Of course, it is ethically impossible for researchers to expose a group of people to something which is suspected to be harmful. But if some of the study population do it voluntarily (e.g. smoking or working with potentially unhealthy chemicals), the researchers may take it as a 'naturally occurring experiment' and simply design a study with an appropriate control group. If the exposure is supposed to be beneficial (e.g. keeping a particular healthy diet), there are usually no ethical problems. It is then more a question of finding subjects who want to participate (which may be difficult enough). A cohort study is often termed a *longitudinal study*. If two interventions are compared in a prospective study, the design is called a (clinical) trail.

In designing a specific research project, researchers will weigh the advantages and disadvantages of each type by what each one requires: cost, time, technical know-how, personnel. A case-control study is usually relatively smaller than a cohort study

and consequently cheaper. This has to do with the prevalence of the phenomenon (e.g. the disease) in question.

Example 3.14

In Denmark the incidence of lung cancer was about 24 in 1962, i.e. 24 out of 100,000 Danes were likely to get the diagnosis during that year. If the sample size calculation of a particular study has shown that it is necessary to compare 100 cases with a control group of 100 or more, then it would work out quite differently in the two kinds of studies. In a case-control study, the sample size would be 200 (100 cases and 100 controls). In an equivalent cohort study 416,000 healthy individuals are initially needed in order to eventually 'harvest' 100 cases of lung cancer during a period of one year. The 100 would serve as the cases and the remaining 415,900 would never get the disease (at least during the time of the study). The smaller the sample selected for a cohort study, the longer the duration in order to identify the sufficient number of cases (i.e. for the cases to accumulate).

One of the weaknesses of case-control studies has to do with the reliability of the recall of past events. If we are seeking data in some old hospital files, chances are they may not have been collected in a standardized way. Different hospitals may have used different techniques to examine lab results, different doctors may assess clinical findings differently, and questions regarding behavior may have been asked in different ways. If data about past events is sought by asking subjects to remember, the risk of recall bias is evident. Exactly how much did you actually smoke ten years ago? How healthy was your diet? How much did you exercise? Did you take any medicine during pregnancy? The answers to questions of this sort are very likely to be uncertain or (even worse) influenced by the present knowledge. For instance women who have given birth to children with malformations are likely to memorize potentially harmful exposures to a larger extent than women who have given birth to healthy children. This is the ghost of *recall bias* coming up again. This risk is minimal in cohort studies, because the 'exposure' data are collected in the present, with the specific purpose of the study and without being influenced by the outcome – the disease - simply because the disease has not occurred yet, but may appear some time in the future.

Defining key concepts

Crucial to the successful design of research is the choice of clear *definitions of key concepts*. Life is full of examples where people use the same word in a number of different meanings often without realising it. It seems important to be clear not only about the definition as such, but also about usage. In fact, we are less concerned about formal definitions than about how concepts are used. Usually it takes different forms in social science and biomedical studies.

Example 3.15

During the past decades, *community participation* has been regarded as an important element of development in the least developed countries and elsewhere. Without going into a discussion of its merits we contend that it has been a politically correct buzz-word that appeared in many policy documents and proposals. But what is it actually? Is it community participating when each household is contributing labour and building materials when a primary school is constructed? Or should they also be involved in the daily running? Is it fair to say that one can only speak of real participation if the community is involved in the priority setting (who says that they wanted a school in the first place?) (Rifkin et al. 1986) What is a community anyway? A well circumscribed village in rural India? Suppose that certain groups are speaking on behalf of the whole village (e.g. the Brahmin caste, the men, the wealthy)? What about the migrant workers who have gone to work in the industry in Hyderabad and are only coming home twice a year? Obviously these issues have to be considered before a meaningful study can be conducted on community participation.

Example 3.15 dealing with community participation shows how social scientists have to be aware of the key concepts that they use in their studies. However, there are equivalent problem for biomedical researchers. When it comes to studies on health issues, there is a need to make *case definitions*.

Example 3.16

Suppose a study wants "to describe childhood diarrhoea among the Sukuma people of western Tanzania". What do the researchers mean by diarrhoea? Is it the presence of 'three or more watery stools within a period of 24 hours'? If so, is it according to the mothers or the children themselves? (who may report differently depending on age). Or should it be reported by the hospital staff from admitted cases in order to get a more clinical and standardized perspective? What is 'watery' anyway? Or should the case definition be based on microbiological verification (the cultivation of various pathogenic bacteria such as Salmonella, etc). Should the dysentery cases (those who have bloody diarrhoea) be included or categorized separately? What about virus infections which are cultivated using different techniques? Or should the study simply be based on somebody (e.g. the mothers) reporting that diarrhoea is present irrespectively of how that conclusion is reached? There are arguments for and against choosing all of the above case definitions, but one thing is sure. It makes a lot of difference for the findings of the study, which one is chosen. This example also shows the complexity of the research process. The case definition which is in focus is directly linked to the choice of data collection methods – e.g. whether to interview children or mothers or go for secondary sources such as hospital files.

Thus, it is important to examine the actual questions asked to generate the data used to construct variables for statistical tests, for many variables that seem straightforward and unambiguous actually turn out not to be so. For example, in many demographic surveys in African countries, we ask about ethnic identity, since we expect some variation along ethnic lines. What is the most appropriate way to ask the question?

Some researchers have asked: What is your tribe? Others, what is your ethnic group? Still others have asked: what is your mother tongue? Or what language did you speak with your mother? Although these questions might produce nearly the same answer in the majority of cases, they are obviously not equivalent.

CHAPTER FOUR

ANALYZING THE DATA

In order to discus data analysis in a meaningful way, we need to remind ourselves of a very basic condition of research. We want to find out something about a 'population'. Typically, a population consists of people, but it can be anything such as goats, cars or birthday parties, i.e. a collection of similar things. As a rule, the population is so numerous that it is impossible to "examine" everybody, whether the examination consists of 'asking a question' or 'measuring a blood pressure'. Consequently, most research projects are forced to examine only part of the population, called a sample. In chapter 3 we discussed how to choose the sample.

Imagine that you have reached a stage where you have collected all the data and are back in the office. The kind of data will of course depend on the data collection methods used. If a questionnaire survey has been conducted, you will have heaps of questionnaire forms. If you have used in-depth interviews, you will presumably have tapes with recordings of the conversations. If the focus has been on nutrition, you are likely to have lists of individuals' ages, gender, height and weight. Generally, you will have data that lend themselves mostly to qualitative or quantitative analysis respectively. Data that contain large numbers of 'units' in the same form lend itself to counting and calculations. It can be the proportion of a group of people answering "yes" to a certain question, the number of patients attending an Out Patient Clinic per day, or the weight of new born babies. Irrespectively of what it is, you can count, calculate, and present the quantitative data in a measurable way. Qualitative data can be in the form of text or images: sentences, video films, notes written by the researcher based on observations, etc. Such data are analysed for their meaning. Sometimes you do some counting on qualitative data, but generally meaning is not derived from counting. Consequently, the two types of data should be analyzed in rather different ways.

Qualitative data analysis

Qualitative data are typically collected with methods that produce text or visual images (see chapter 3). The text is typically a document containing a transcribed interview. The pictures can be a video recording of an event or a child's drawing. Of course some aspects of the data can be expressed in figures. For instance, we can count how many times a certain word is used by an interviewee, or how many times two people are shaking hands in a video recording. But these calculations describe only isolated phenomena out of a much more complex total representation and are not meaningful in their own right. Thus, qualitative data have to be analyzed in a way

appropriate to the form of the data collected. Without going into details, we will try to give a general idea about some of the elements and principles of such an analysis. We will concentrate on textual data.

Classifying 'knowledge'

Social science studies often seek to explore the 'knowledge' (perceptions, beliefs) of a population regarding health problems. Usually such studies are based on '*what* questions'. This raises a number of issues that we need to address. What kind of 'knowledge' is it? Is it the 'scientific truth' (e.g. the biomedical explanation of the cause of a disease)? Or is it local knowledge, the way members of a certain society see it? It makes a lot of difference which of these two perspectives forms the basis of the research, particularly in research on health and illness. Are we interested in knowledge that fits with the biomedical view of what is relevant, and gear our research questions to those principles? Or are we interested in local knowledge of illness and medicine? Anthropologists often distinguish between '*emic*' and '*etic*' perspectives. An emic perspective refers to local knowledge, the concepts and knowledge of a specific population, the view from the inside. An etic view, on the other hand, is one from the outside (Pike 1971). The view from the outside may be a biomedical one, but not necessarily.

Example 4.1

> Among many ethnic groups all over the world, sorcery is perceived to be a fact of life and is seen as an important cause for misfortune – be that illness, bad harvest, or lack of success in education. Suppose that a study revealed this to be the case. An 'etic approach' would discard this explanation as supernatural and unreal, and look for other causes. An 'emic approach' would make it a point to make a thorough description of what people know about sorcery and how people responded to it without judging whether it is 'true' or not.

We are not saying that one perspective is better than another. But we want to make two points. The first point is that it is important to realize which approach is the most appropriate for the program that the commissioning agency wants to implement.

An example is given below.

Example 4.2

> Health projects aim at controlling (reducing, eliminating, eradicating) diseases and the interventions usually involve an attempt to modify peoples' behavior by the means of health education. More and more implementers realize that health education programs that aim at introducing new biomedical knowledge and change practices need to take the target population's present perceptions and practices as the starting point. To simply try to teach biomedical (etic) knowledge is bound to fail. Consequently, many studies are conducted to describe the emic perspective in order to give the implementers a better basis for action. (See also example 2.3)

The second point is that it is important to be able to recognize research that confuses the two approaches as described in example 4.3.

Example 4.3

> Once we met a student who set out to conduct a study on health seeking behavior at the household level in a rural part of an African country. His objective was to get a picture of the way decisions were made at the household level. Would the caretaker take the sick child to the government clinic or to the traditional healer for treatment of fever? Would he/she simply go and collect some herbs, which were considered potentially curative? We know that most people in developing countries (and elsewhere) do not think of illness in biomedical terms. As they are locally diagnosed, illnesses are usually not equivalent to biomedical diseases. Hence a researcher should not inquire where the children's caretakers seek treatment for 'filariasis' or 'German measles' (i.e. the etic terms). More often than not this will not make sense to the informants. Instead the researcher should explore the local illness entities (the emic perspective) and subsequently find out how they are treated. The mistake that this particular student made was to apply the biomedical terms when interviewing the study population about their health seeking behavior. By asking etic questions and using etic concepts as the units of analysis, he could simply not do what he claimed he was doing.

If a study sets out to understand the emic perspective, as seen in the perceptions of diarrhea among Swahili speakers in Lubumbashi in the former Zaire (as described in example 4.4 below), the data analysis will be a process of understanding how individuals make their own diagnosis of illness. The end result will be an overview of the relevant categories and their distinctive features. Such research is similar in some ways to what is called componential analysis, ethnomedicine or ethno-science. Whatever we call it, in such a context we study local knowledge of a particular phenomenon.

Example 4.4

> Some years ago we conducted a study in Lubumbashi, Zaire of how Swahili speaking mothers diagnose different types of diarrheal disorders (Yoder, 1995). We organized some 40 discussion groups (3 to 5 persons per group) with women to ask about their knowledge of childhood illness. In Swahili we asked what childhood illnesses they knew about, and then selected certain ones to ask about how they recognized those illnesses. Doing things this way makes it possible to obtain information without suggesting any kind of answers at all. We also asked about the treatments that were possible or appropriate for the illnesses we were focusing on. We found that these mothers recognized five different illnesses that we would classify as diarrhea, with each one having a separate set of symptoms and recommended treatments.

The results of the research described in example 4.4 gave us a complete picture of the illnesses that mothers diagnose in Swahili, complete with the symptoms associated with each illness. Some researchers call this a *folk taxonomy* to distinguish it from the biomedical taxonomy.

Analyzing observations

The analytical construction of folk taxonomies can reveal the structure of concepts and categories meaningful to the local population, but it says nothing about what people actually do. As mentioned in chapter 3, the best way to study people's practices is to observe what they do, rather than just ask them. Observations of events by the researcher as expressed in field notes, or more structured observations based on guidelines given to field assistants, are common forms of such evidence.

Example 4.5

> We once studied the medical practice of a traditional healer in southern Zaire, a man who treated a wide variety of illnesses, both mental and physical. I wanted to find out how he prepared and administered his medicines. So I visited him every day to accompany him into the bush to collect plants and other objects and ingredients for his medicines. I also assisted in his preparations and in giving medicine to his patients. I kept a careful written record of the process of finding and collecting ingredients, preparing medicines, and administering them to his patients. The analysis of these data consisted of identifying key terms and concepts, seeing how elements occur in a sequence of events, and considering how the elements are logically related to each other.

Identifying key themes

The analyst should go through the data (texts or graphic material), select certain issues or elements of special interest, and then mark those sections. This marking process is called *coding*. Each time a conspicuous issue appears in the data, the

researcher ascribes a note indicating which issue it is. The researcher decides how the coding is to be done. The units to be marked may be key terms (e.g. the local names of illnesses) or an event (e.g. accounts of domestic violence) appearing in the interviews.

Once the coding has been completed, the researcher can then extract all the relevant pieces of text that deal with that particular issue. The coding is often done in several rounds as the researcher finds different ways of marking units within the text. Coding can be done manually by writing on paper, or by using a computer software program. Coding is also part of quantitative questionnaires surveys. The predefined, closed answers are each given a code (usually a number) which allows them to be entered into a software program for quantitative analysis. However, this form of coding is rather different from that of qualitative analysis which is a much more creative and open-ended process.

Looking for links

The next step of qualitative data analysis could be to look for underlying patterns of co-variation and causality. We can also check to see if there are some kinds of associations between the various issues. A way to achieve this is to organize the data in matrices or tables. A given matrix focuses on a set of issues, which are set up against each other. By visualizing the co-variations like this, patterns will appear.

Example 4.6

> We recently designed a study of complementary feeding of infants in the Mopti region of Mali (see examples 2.1. and 2.2). Given the high rates of stunting (malnutrition that leads to low height for age) among children three to twelve months old, we wanted to find out if the way children from four to twelve months old were being fed could explain the stunting. We observed and interviewed mothers whose children were well nourished and those whose children were poorly nourished. We then organized data into tables comparing the well nourished with the poorly nourished, looking for differences in frequency and quality of feeding. We found little difference in the frequency of feeding, but that the more poorly nourished children were fed a wider variety of foods.

Researchers do not agree on the extent to which analyses of meaning and social context (qualitative analysis) data can shed light on causal relationships, and this is not the place to go into details. Some argue that only statistical proof of association can identify causal relationships. However, qualitative data presented in tables and matrices can give strong indications of such links. This points to one of the important functions of open-ended studies based on qualitative data – they can be '*hypothesis generating*'. In other words, such a study can give raise to interesting ideas about how things are connected. If more firm (and quantitative) confirmation of such hypotheses

is desired, it is subsequently often possible to design another kind of study to test such a hypothesis (asking a *what* question). The point is that the more open approach of social science studies using qualitative methods (and especially those asking '*how* questions') are more likely to discover new and novel connections, than are the more structured approaches.

A qualitative approach to analysis may set up *models* that encompass and explain a socio-cultural phenomenon. A model is constructed by the analyst by moving back and forth between the data and the model, which is modified in the process. The extreme cases are held up against the model, which is then changed in order to encompass the extreme cases (see 'outliers' below under 'quantitative data analysis'). Similarly, the search for the so called 'negative evidence' can contribute to a shaping of model, which grasps the socio-cultural reality as loyally as possible. Whereas quantitative analysis (see below) goes for average measures of a phenomenon rather than the extremes, qualitative analysis considers the extreme cases as well. Within this context, a model is not either true or false. Instead, it is judged by its explanatory power, or goodness of fit. The more clearly the model helps to explain a given social or cultural phenomenon, the better the model is, whether it covers all elements or not. This contrasts with models within natural science, where just one opposite case can falsify the model completely.

Example 4.7

> UNICEF created a model to reflect the relationships it sees as operative in the determination of the nutritional status of young children. That model has been slightly changed and is being used by Macro International in their nutrition studies.

Consider variation and context

All human beings are somewhat different from one another. Though we may have the same cultural roots, come from the same social group, and believe in the same religion, we still think and act differently. There are many explanations possible if we choose to seek them. However, in qualitative research, we examine individual variation. It is the hallmark of qualitative data analysis that attention is paid to this variation; that it is kept in focus and not leveled out as in quantitative analysis when calculating a mean. It is not that qualitative analysis is not looking for general trends. It does. But it should be done in a way that respects both history and social context.

However, life is even more complex than that. On top of the variation between individuals, to which anthropologists pay a lot of attention, there is also the many and constant changes in the 'environment'. Thus, another characteristic of qualitative data analysis is the close attention it pays to context. Where most quantitative research seeks to hold the context 'constant', and thus neutral, in order to pursue a research

question, qualitative analysis makes it a point to integrate the context into the analysis. There are always particular circumstances and exceptions to the rule. Decisions are made based not only on the actors involved in the transaction, but also on the characteristics of the setting in which it takes place, i.e., time of year, age, gender, educational level of other persons present. A good qualitative analysis brings this kaleidoscopic perspective to the fore.

Comparison

One of the basic characteristics of most research is its concern for comparison. New insights can be gained by comparing groups.

Example 4.8

> Suppose we begin with wondering why teenagers in Baltimore engage in unprotected sex. Do the teens themselves share our conception of unprotected sex? Do we assume that they do so because they do not understand the risks they are taking? A study of adolescent sexuality in Accra and Kumasi in Ghana found that some female adolescents (12 - 20 year olds) engage in sex for monetary gain (Nabila & Fayorsey 1996). Should that be one of our hypotheses for Baltimore? Are there social situations that facilitate casual sex? Are parents not paying attention to what their children are doing? Are these young people exposed to models of behavior that promote risky behavior? What can be learned from the comparison between the two populations?

In the presentation of the results of qualitative analysis, direct quotes play an important role. Brief extracts of interviews are cited in reports or articles in order to illustrate the points made by the researcher in the analysis. Finding just the right way to summarize the contents of 50 or 100 interviews that reflects both the variety of perspectives expressed as well and the dominant points of view (key themes) is an art learned over time.

Quantitative data analysis

Quantitative data analysis is very much a question of statistics. Statistical methods seek to describe and analyze populations taking variation into account. In chapter 3 we discussed "random error". One can say that random error is the element of uncertainty that natural variation induces into research findings. As Campbell and Machin (1990: 2-3) have expressed it: "Measurements on human subjects rarely give exactly the same results from one occasion to the next... These differences may be due to real effects, random variation or both. It is the job of the analyst to decide how much variation should be ascribed to chance, so that any remaining variation can be assumed to be due to a real effect. This is the art of statistics". Example 3.8 where a coin was tossed illustrates this clearly.

Statistics involves lots of formulas and calculations, and we realize that many of our readers do not have the training or the need to master this. We will therefore try to deal with statistics using a minimum of figures. We will concentrate on the logic behind the symbols and try to clarify what they mean.

Example 4.9

> Let's go back to example 3.1 where we measured the height of young men in Chicago. Let's say that we have now done the study and measured the height of eleven individuals: 138 cm, 165 cm, 175 cm, 178 cm, 178 cm, 182 cm, 184 cm, 186 cm, 189 cm, 191 cm, 192 cm. These are our 'raw data'. Looking at them we get a pretty good idea about the situation. But had we made 300 measurements, it would have been much more difficult to get an overview.

There are two types of statistical analysis: *descriptive analysis* and *inferential analysis*. Descriptive statistics deals with ordering the data within the sample, whereas inferential analysis has to do with the way conclusions are made about the population based on the data collected from the sample.

Descriptive statistics

Descriptive statistics are used for organizing the quantitative measurements of the sample. It seems fairly straightforward.

One of the simplest ways of organizing data is to calculate the *'mean'*. This is done by simply adding the measurements and divide by the total number of cases. In example 4.9 the average height is 178 cm. Let's look a bit closer at the figures. It seems that one of the young men is very short at 138 cm. Provided that it has been measured and entered been correctly, he is presumably a dwarf, which does not disqualify him for the study. However, as there are only eleven measurements, his height is pulling the mean downward in a disproportionate way. Technically, the height of 138 cm is said to be an *outlier* – an extreme and atypical observation. In order to compensate for that, the *'median'* is sometimes used instead of the mean. The median is the value of the middle observation, in this case that of number 6 out of the 11 (182 cm). It is not that the mean is truer than the median or vice versa. It is simply a question of presenting actual data in different ways in order to give us an accurate reading of "where the middle is".

However, in descriptive statistics, finding the middle is not enough.

Example 4.10

> In example 4.9 we found that the mean was 178 cm. However, the same mean could have appeared in many different ways. It could have been based on the measurements: 153, 158 cm, 163 cm, 168 cm, 173 cm, 178 cm, 183 cm, 188 cm, 193 cm, 198 cm, 203cm. Or it could have been eleven measurements of men with exactly the same height, 178 cm.

Obviously, we need an indication of how much the observations are spread out. The easy way is to show the *range*, i.e. the two extremes between which all are situated. In the first part of example 4.10, the range is between 153 cm and 203 cm, whereas in the last part there is actually no range. It gives us a pretty good idea about how concentrated the measurements are around the mean, but it is also sensitive to outliers. In example 4.9 the extreme case of 138 cm influences the range significantly (138 cm to 192 cm). Another way of showing how far the observations are spread out is by using the *standard deviation* (SD). The mathematics is slightly more complex here, but basically a formula is used whereby the difference between all the measurements and the mean is calculated and condensed into one figure, the SD. The larger the SD, the more the observations are spread out. In example 4.10, the SD is about 17 cm. This gives an indication of how closely the various measures are grouped around the mean. Both the mean and the SD have the same appellation as the measurements themselves, in this case centimeters.

Inferential statistics

Inferential statistics are slightly more complicated to grasp. This is where we try to draw conclusions about the whole population based on the findings from a sample. Making this jump from sample to population is influenced by many factors. One group of factors has to do with the way the study is designed, i.e. whether there is some sort of selection bias or information bias (see chapter 3). However, in this section we will concentrate on the factors that have to do with random variation.

Basically, inferential statistics is build on the thinking that there is a "true value" (a *parameter*) in the population that we are trying to find. However, due to random variation (error) we cannot quite get there. We need to make an *estimate* based on the measurements in the sample. The better the study, the closer the estimate gets to the parameter. The estimate may even hit the parameter exactly, but we will never know. What inferential statistics can do for us is to describe approximately how much we can rely on our estimates.

One way of indicating the precision of an estimate is by using a confidence interval (CI). Again we will skip the calculations. Suffice it to say that a calculation is made (based on the sample size and the SD) which gives an interval within which the 'true value' of the mean (the parameter) is probably lying. The CI (95%) in the first part of example 4.10 is (166.86 cm ; 189,14 cm). That means that the parameter is situated within this interval with 95% certainty. Or in other words, that there is a 5% risk that it is outside. If we want to have a higher degree of certainty (e.g. 99%), we have to pay a price. The CI (99%) for example 4.10 is (162.15 cm ; 193.85 cm). As it can be seen it is broader. On the other hand the CI (90%) (168.94 cm ; 187.06 cm) is more narrow. The high degree of certainty in the statement is compensated by a broad safety margin in the CI and vice versa.

Example 4.11

In example 4.10 the CI (95%) of (166.86 cm ; 189.14 cm) was based on a calculation which involved a sample size of 11 (i.e. there were 11 measurements of height). Let's imagine that the study involved 100 measurements instead and that it resulted in exactly the same mean (178 cm). We will not waste your time with lining up all the 100 measurements, but have simply invented some values that all fall within the range of 138 cm and 192 cm. Based on this larger set of data, the new CI (95%) is (174.71 cm ; 181.29 cm). As can be seen, it is much narrower. (The calculations are slightly simplified for the purpose of clarity).

In example 4.11 the confidence intervals indicate the level of uncertainty of a mean in one population. However, quite often a study aims a comparing two populations in one way or the other. This is illustrated in the following example.

Example 4.12

Let's suppose that we are conducting a study which aims at comparing the height of young American men (18-25 years of age) in Chicago with young men of the same age of Maasai origin in Eastern Africa. Our data have shown a mean in Chicago of 178 cm and among the Maasai of 184 cm. The analytical question is now whether this is a "true" difference? Are the parameters of height really different in the two populations? The Maasai are known to be very tall, so we would be inclined to believe that the difference is real. But do our data allow us to conclude that they are?

In order to address the question in example 4.12, researchers calculate the level of significance, which is usually shown as a figure less than 1. If the level of significance is 0.05 (or as some says 5%) for the difference in height, it means that there is a 95% certainty that this is a "true" difference (or a risk of 5% that it is a random variation). If the situation occurs that a study indicates a significant difference between the two

estimates though in actual fact the two parameters are not different, the researchers are misled.

However, it could also be the other way around. Sometimes two parameters are actually different, but the study fails to show it. Such a situation is much more common than the situation described above and should be carefully considered not only during the analysis phase, but already during the planning phase (see chapter 3). In order to understand this, we need to remind ourselves about how a confidence interval (CI) calculation is based on the sample size and the SD, because the level of significance calculated in an equivalent way.

Let's take the sample size first and still take example 4.12 regarding the height difference between the Chicago and the Maasai males as the reference. We have shown above that the larger the sample size; the more certain we are about the outcome. In the example there was a difference between the means 178 cm and 184 cm. Based on common sense we would be more reluctant to trust this to be a true difference if only a few individuals had been measured in each place. Only one extreme case (outlier) like the dwarf of 138 cm in Chicago would pull the mean downwards. However, if the two means were based on 1000 measurements in either population, a single outlier would not change much. This common sense approach shows how important sample size considerations are. Fortunately, it is actually confirmed by the formula. There is another factor at play here, however – the size of the difference.

Example 4.13

> In example 4.12 the mean height of the Chicago males was 178 cm whereas that of the Maasai men was 184 cm. Suppose that our study had shown the mean for the Maasais was 198 cm instead. In other words the difference between the two estimates was 20 cm rather than 6 cm. If we were assured that the study had been conducted according to all the best standards, we would be more inclined to believe that this reflected a true difference between the two parameters. Simply because, everything being equal, it would be less probable that it was caused by random variation and a single outlier or two.

Thus, we have a situation where the certainty of a study finding is influenced by two factors: the sample size (the bigger the size the more certainty about the outcome) and the size of the difference (the bigger the difference the more certainty). Or to put it in other words: if the actual difference that we try to show is big we don't need a large sample size, but if we expect it to be small we need a big sample. We are not in control of the magnitude of the difference, which is based on parameters in the "real world". But we do control the sample size. In order to ensure that the study will bring out the certainty that we look for, we should make the sample size large enough. If we

don't, we risk making a mistake. Technically, one says that the study was planned in such a way that it didn't have enough *power* to come up with an answer. If this is not taken seriously, the resources spend on a study, which is well thought out in every other way, may risk to be wasted simply because the study doesn't have enough power. Millions and millions of dollars have been spent on studies that ended up as inconclusive because of low power, since the proper sample size calculations were not made.

In this explanation we have used the example of a difference between two means. It could have been many other things, e.g. the difference between two ratios, relative risks etc. There is a number of different formulas to calculate the sufficient sample sizes, which are beyond the present book to explain. It should also be emphasized that this issue is only relevant for studies that base themselves on quantitative data.

There are a number of statistical tests which can be applied. Again we realize that the readers of this book don't need to go deeply into this. Managers don't need to know the formula and how to use them. All that is needed is to know that there are many different statistical tests with a lot of odd names (e.g. Student's T-test, McNemar Test, etc.) and that the choice of which one to use depends on a lot of preconditions. For instance it may depend on the size of the sample or whether the variable is discrete (like the number of babies born) or continuous (like the measure of height). The critical question that the manager should ask is to request the consultant to explain in plain words why this particular formula/test is being used. An answer that "the statistical software program SPSS has been used" is not sufficient. It is like saying that "a car has been repaired in a garage". What we want to ascertain is that the mechanic can justify why he "changed the piston rings and not the spark plugs".

CHAPTER FIVE

REACHING CONCLUSIONS FROM THE DATA

No matter what our research strategy is, we test our conclusions by asking a series of questions of the data and the conclusions. Do we have different types of evidence that all point in the same direction or do they point in opposite directions? Are there alternative interpretations that should be considered? Are there major portions of the data that were not used in reaching our conclusions? How do we evaluate the strength of a particular conclusion? Can we generalize to other populations?

In both qualitative and quantitative approaches, the researcher may have a hard time reaching firm conclusions. However, researchers should always be clear about the research questions, the assumptions, and the methods. No matter what the results are, if a program manager can follow the logic of a research project, he or she has a basis to draw on to decide if the results are of any use to the project.

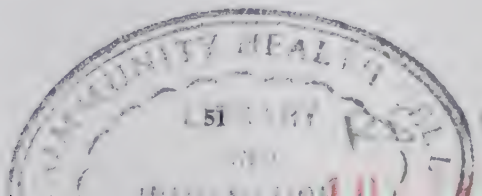
Did we get what we were seeking?

At this stage it is important to compare the conclusions with the original research objectives. Did we get answers to the questions raised – fully or partly? Have all the specific objectives been addressed? If not there may be good reasons. Maybe it was not practically feasible. Maybe the researchers were uncovering new insights that forced them to change the study mid-stream and move in a new direction. But they may simply have forgotten. In that case the program director should encourage them to go back and provide the answers. If they have the data they can do it. If they don't, it may be impossible and the study may be a failure.

But there will also be cases where we get answers to questions that we didn't even ask. The researchers may have come across valuable insights which they picked up on the way.

Example 5.1

In example 2.9 a study was designed to estimate worm infection prevalence among primary school children to pave the way for a school health program. The project was based on the assumption that children are accessible in school for the study. Suppose that the researchers had come back and told the NGO that this was not the case (because of low enrollment), and that the basic idea of the planned intervention (a school-based de-worming programme) was dubious? That was not part of what the researchers were sent out to do, but such a finding would indeed do the agency an unexpected favor.



Are the findings valid?

This is a bit complicated. It is a question of estimating how the researchers have dealt with the random and systematic errors as described in chapters 3 and 4.

Random errors are of particular interest in studies using quantitative methods. Statistics is the way of dealing with it. Or rather, the way of giving a quantitative assessment of the uncertainty of a given statement. This is expressed by either a confidence interval or a significance figure. Depending on how strong the design was, the power of the study will allow more or less firm conclusions.

Systematic errors are more complicated and more relevant to studies using both qualitative and quantitative data collection methods. Are there any indications of selection bias, which would be a reflection of inadequate sampling? Systematic errors can also be caused by the data collection methods themselves. Have the biomedical measurement instruments (e.g. the thermometer) been well calibrated? In social science *triangulation* is used as a means of reducing the bias involved in data collection. Triangulation has three aspects. Data should be derived from more than one method. Data should ideally be collected by more that one fieldworker and more than one group of people should be studied. The more all these different sources point in the same direction, the stronger are the conclusions.

Example 5.2

Suppose we have conducted a study of how young people find sexual partners in Brazzaville in order to provide information for an HIV prevention program. We have data from observations on how young people interact in church associations, in schools, in bars, and in youth clubs. We have accounts of the experiences of 50 young people from two neighborhoods in forming friendships, spending time together, and eventually having sex, as well as accounts of more casual encounters. The data are in the form of histories of these individual experiences. We also have data from interviews of persons who direct activities for youth. The individual accounts provide the concepts and the vocabulary used to describe social interaction in this population, accounts that can be compared with the observations conducted. The combination of observations and accounts make it possible to describe actions that often lead to risky sex.

This research in Brazzaville illustrates the kind of study that qualitative research does well. The research focuses on actual social interactions that often occur and so that can be observed, and it solicits individuals' accounts of their own experiences in these same kinds of interactions. In this way we begin to understand how young people in Brazzaville go about having sex.

Within data analysis there is also a potential for systematic error. Regarding qualitative data, a good study should indicate the steps in the analysis including the development of models and how extreme cases (in the statistical jargon 'outliers') and contradictory evidence have been dealt with. In quantitative data analysis (which is basically inferential statistics) there are numerous possibilities for misinterpretation. However, this is beyond the scope of this book.

Example 5.3

> We conducted a baseline and a follow-up survey in two provinces in northern Zambia in 1991 and 1992 with a representative sample of adult men and adult women in order to evaluate the impact of a radio soap opera about AIDS (called *Nshilakamona*) which was broadcasted in Bemba, the language most people spoke in those provinces. The proportion of persons who said AIDS was the most serious health problem in their neighborhood increased from 7% in the first survey to 23% in the second. How might this be explained? We considered two main possibilities: program effects, or more AIDS cases. Since the change occurred in both populations who listened to the AIDS program and those who did not, the change was not an effect of the radio program. It may have been due to the increased number of people with AIDS who were visible in the neighborhood. That is possible but we do not have any real evidence for that. We were not able to think of other explanations to consider. (Yoder & Hornik 1996)

Reports or articles of studies using statistical analyses to measure the significance of results routinely report on the strength of the associations found and on the construction of the variables.

To which extent can the findings be generalized?

The possibilities of generalizing findings to other populations depend on the study design and on the elements examined in the study. On the one hand, studies which base themselves on random sampling of rather large study populations are usually strong in this aspect. On the other hand, more ethnographic studies using qualitative methods and purposive sampling generally find it harder to argue that their findings apply to a broader area or population. A result explaining the influence of vitamin A on infant mortality is likely to be global in the sense that it is valid in Venezuela as well as in the Philippines. The potential for generalizing about the findings on the prevalence of birth complications among Kikuyu women in Kenya depends on the sampling procedure, i.e. the sampling frame, sampling size and research procedures. If the researchers have conducted an ethnographic study in a particular village it is uncertain to which extent the findings can be generalized. It will depend on the issues in focus and the way the researchers make their case. But most likely the study findings will only apply to the local area.

Checking the assumptions

Some researchers claim that assumptions are not relevant for exploratory research (which simply aims at 'finding out what is there'). However, experience tells us that we all have assumptions about the phenomena we examine even if our actual knowledge at the beginning may be limited. We still act on our assumptions. Articulating them is always an illuminating exercise. This is the time to revisit the assumption made at the planning phase and see whether they have been taken into consideration in an appropriate way. Or were new assumptions uncovered in the process and were they dealt with sufficiently? Most research designs in qualitative research do not discuss their assumptions, although we should all follow that routine.

Even in cases where the researchers are not explicitly stating their assumptions, clues may be garnered from an examination of the questioning guides. The questions asked (and not asked!) point to the areas where the answers are supposed to lie. As suggested earlier, the quality of qualitative research can be judged by the degree of transparency in the logic of the design, and the links made between the data presented and the actual findings.

Are the findings useful?

Presumably most program directors commission research to achieve concrete results. We can think of at least three. First, large agencies may be involved in policy making or advocacy and may need research evidence to direct and support them.

Example 5.4

> One of us has been involved in a child research in western Kenya for quite a few years. Since 1998 we have focused on orphans. Our research revealed that one out of three children under 18 years had lost at least one parent and one out of nine had lost both – presumably mainly due to the rampant HIV/AIDS epidemic in this area. Such figures call for major interventions. This research was not commissioned by any agency, but we have tried to spread the news to local and international policy makers (Nyambedha et al. 2001).

Second, some agencies will have to make choices regarding which health problem to address (e.g. provision of better ANC or school health?) in a particular area. Or they may need guidance on where to work (e.g. in which district of Tamil Nadu or among which castes?). A descriptive study can provide the answers. This was outlined in example 3.11.

Third, an agency may already have decided on the area and the target group, but needs evidence for how to best conduct the intervention. The examples of complementary

feeding in Mali (examples 2.1 and 2.2.) and the health education on malaria (example 2.3) are such cases. This is also seen from the following example.

Example 5.5

A study of health seeking behavior in households with children under five years of age in the Maninka speaking areas of Upper Guinea. The study has two objectives: one, to determine what signs of illness parents and other care-takers recognize in their young children, and two, to discover how care-takers have responded to those signs of illness in order to restore health. Interviewers visited randomly selected households over a period of three months every other day, observing what has occurred and obtaining accounts of who participated in the response to illness. This resulted in some 160 case studies of response to illness. The case studies show what options care-takers have in their response, and the kinds of elements they consider as relevant in making their decisions. Assuming the villages chosen for the study are typical of Maninka speakers, the findings will allow us to suggest ways that the health service could intervene to improve local options for treatment, and to generalize about the patterns of care given to such children in the region.

Did any new research questions occur?

Good research is characterized not only by giving appropriate and reliable answers but also raising a number of new questions. Some of these ideas may be of general interest and can be published in scientific journals. Others may be of direct relevance to the program. Let's go back to the school health pre-study again (example 2.9). Suppose that researchers didn't find the actual proportion of children of primary school-going age not being enrolled. Suppose that they simply came back and pointed out the issue. Then the program director immediately had to commission new studies to find out the age distribution of worm infections among children of different age groups, the proportion of absenteeism and the school enrolment rate.

CHAPTER SIX

PRESENTING THE CONCLUSIONS IN THE APPROPRIATE FORM

The Audience

The results of an applied research project have different audiences. These are the sponsors of the project (program directors), researchers with similar interests to the researchers, and the study population. Each of these groups has a different vested interest in the results of the research and the authors should address them in a different ways. The sponsors (readers of this book) expect a full technical report with an executive summary stating the main findings. They are accustomed to reading technical language that is simply stated.

The sponsors determine the form of the final report, but the researchers are responsible for the technical quality of the text they deliver. Good technical reports present at least four aspects of the project: 1) the objectives and assumptions; 2) key parts of the evidence, the data, are presented in ways that show how they were collected; 3) the findings are presented in ways that show the evidence and how it was interpreted; 4) the implications of the findings for the program are laid out. Far too often these technical reports blur the boundary between data collected and interpretation. The reader must be able to follow the research process and judge for him or herself if the findings are convincing. As a special point in relation to the sponsors, the contractors have to make a financial report (e.g. an audited account) in one form or another as well. The sponsor and the research team should try to be in continuous dialogue regarding the research process in order to prevent misunderstandings.

Some researchers may not be able to make clear policy recommendations from their research. It takes some practice to shed the habit of safe-guarding everything you say with reservations and venturing into coming up with solid recommendations. Likewise it takes a bit of practice to express research issues in a relatively simple language with a minimum of professional jargon. Some will claim that it is impossible. We contend that it is perfectly possible with a little effort.

Apart from delivering a report to the sponsor, researchers are usually also keen to publish articles in peer-reviewed journals. However, there is a potential conflict here. Some sponsors, like readers targeted by this book, may feel that they paid for the study and that they consequently 'own' the results. The sponsor may not wish the data to be published maybe because the findings are inconvenient or embarrassing. But suppose that another NGO wants to do similar things some years later. They can learn from the experiences. Maybe they even want to operate in the same country? If the

findings are only available as one of the innumerable reports in the vast realm of gray literature, it is so much more difficult to obtain. Consider how grateful you would have been to get hold of a similar report in the planning phase. Besides we find it unethical to repeat a study if unnecessary. It is unethical to spend the money and unethical to waste the time of the study population. Finally, it may actually contribute to the good image of the sponsor to be seen as a promoter of scientific publications.

The study population should expect presentations of the main findings in simple and direct language applied to their situation. The venue for such presentations will be group meetings with ample opportunity for community members to react to the findings. If returning to the study population with the results is not part of the original design, this will not happen. This aspect of reporting carries with it two dangers not present for the other forms of reporting: we risk exposing our research as being of little relevance, and we risk misunderstanding on the part of community members. If community representatives have been involved (maybe as part of participatory research projects), then it is likely that the research is not completely misguided and that they will have insightful comments and suggestions once it has been completed.

Example 6.1

A few years ago one of us designed and directed a study of knowledge of childhood illnesses and the risk of contracting intestinal parasites in the school environment in two languages (Susu and Fulfulde) in Guinea. Since one of the objectives of the research was to enlist the assistance of parents in improving the school environment, we planned to present the results to parents and ask what they thought they could do to help the schools (assure the school of a safe water supply, build or improve latrines, counsel the food venders about cleanliness, etc.). In the first place where we returned (Susu speaking territory) to report on our research, the first reaction of the parents was amazement that we actually returned. Though they had been the subjects of research of various kinds before, never had the research team bothered to return for a discussion of the results.

Sharing results with the study population will usually take the form of dialogue meetings rather than written reports - community meetings, group conversations, etc.

Example 6.2

> A research project in East Africa operated in primary schools and studied the interface between health and education. One day the mother of one of the pupils called me, and expressed a big concern about the number of drop-outs from the local school. "Couldn't we look into this issue as researchers?" So we did. This conversation has led to two major studies on that issue. The point we try to make is that giving feedback to the community members is not only a 'tedious duty'. Establishing an open dialogue with the target population can serve as a source of inspiration helping you to get new angles on the data analysis and identify new, important research questions. This issue is further elaborated in chapter 8, where the issue of participation is mentioned.

The community feedback is usually neglected. It is not budgeted for and it is considered too laborious. "Who cares anyway, now that we got our data?" Apart from the ethical aspects of such a stand, there is a more practical one as well. A lot of studies have been conducted in most parts of the world and if the previous research team were misbehaving in one way or the other (usually by not keeping promises) the next team will have a difficult time. If the group sponsoring the research (e.g. the NGO) is expecting to operate in an area for a relatively long time (e.g. several years), which is the most sensible way of contributing to development anyway, the importance of giving community feedback becomes especially important. It will be in the best interest of the agency itself.

The format

Usually, the study will result in a report and become part of the gray literature that is never published and therefore not widely accessible. Occasionally an article may also be published in a scientific journal. Volumes of textbooks have been written on how to write reports in the most appropriate way and we are not going to repeat that here. We mention just a few general issues below.

First and foremost, the format should be adapted to the audience. One writes differently for sponsors, researchers and the study population (unless you want to talk to them instead). This has a bearing on terminology, size and main emphasis. The sponsors will presumably be interested in many details and researchers more in the methodology and theoretical framework. The study population will be interested in concrete recommendations that can be put to immediate and practical use.

But there are other criteria related to quality which should always be considered. The language should be kept as simple as possible, even in scientific articles. The meaning of the key terms and all the abbreviations should be clear. The various elements (e.g. data collection methods or recommendations) should be presented in a separate section and in a logical sequence. Finally it should not be too long! Researchers have

a tendency to find every little detail very interesting and very difficult to do without. A voluminous report is quite often an indication that the researchers have not worked enough with the material and distilled the essentials. A report containing loads of data (e.g. presented as dozens of tables) without any attempt to synthesize, explain and compare with findings of other studies is clearly a danger sign. Furthermore, all reports should contain policy recommendations, e.g. "the most important health problem in the area is the respiratory infections due to the pollution from the nearby factory and this should be addressed immediately". Finally, all reports should have an executive summary, which include three issues: What happened?; which were the findings?; and what is the meaning of the findings?

CHAPTER 7

ARRIVING AT THE RIGHT PRICE

A detailed budget should be seen as an integral part of a project proposal. Although it is usually difficult to predict just how each research activity will go and how many days will be spent, we still need to plan all activities ahead of time and calculate the budget accordingly. Once the data collection activities have been described, and the time needed for literature search, data analysis and writing up has been decided, we can make our budget.

A budget should be both fair and reasonable. On the one hand, the consultant or the contracting agency should not get an extravagant profit. What profit margin seems reasonable depends on what people with similar contracts have been paying. Thus we need to find out what the precedents are, but then pay an amount that the consultant considers fair. On the other hand, signing a contract that does not provide for sufficient funds to complete the job properly creates problems for everyone. Thus we must be sure not to under-fund our research either.

Building the budget

The easiest way to put together a budget is to divide the process into small steps which we describe below.

Step one – identifying the budget lines

The first step in making a budget is to make a list of all the possible expenses you expect to have, and arrange them into categories. Budgets for field studies normally have the following types of expenses:

1) Salaries: pay for labor of all kinds - supervisors, interviewers, observers, drivers, translators, data entry clerks, support staff, etc. Salaries are most often calculated by the day for shorter projects, by the month for longer ones. We should expect to pay somewhat more than what people are earning working for the government, and as much as the private sector pays.

2) Per diem expenses: individuals receive funds whenever they travel away from home to pay for food and lodging.

3) Transportation: local and international travel. While the latter may be easy to estimate over the phone, the former requires a detailed plan with the number of persons and vehicles specified. The local consultant may need to rent several

vehicles along with drivers, or he may be able to obtain vehicles on loan from one ministry or another. In the latter case, the project will need to provide fuel and other services items. It should be borne in mind that local transport not necessarily identical to vehicles. In some instances bicycles and other low cost solutions are quite often more appropriate.

4) Equipment such as tape recorders, computers, printers, perhaps flashlights.

5) Consumables such as paper, pens, pencils, tapes, diskettes.

6) Communication: funds for telephone, fax, postage, DHL, E-mail services.

7) Dissemination costs: printing and distributing the report. Funds will be needed to present the results to the population involved, and to interested scholars and policy makers.

8) Miscellaneous: funds for the unexpected. This can occur at two levels. At the project level there may be unexpected extra expenses, e.g. drastic increases in fuel prices or extra salary expenses due to delays in data collection. At the individual level the research team will inadvertently face situations in which their direct assistance is warranted. Examples: an orphan with malaria and nobody to cater for his treatment, or the provision of a water tank to a school which has been especially helpful during the project implementation (see also ethical aspects in chapter 8).

Field studies typically incur expenses in each of these categories. However, some contractors or sponsors have their own formats for budgeting, forms that combine these categories differently. For example, the categories of communication and dissemination may be put into a category called Administration. Often there is an *overhead* to calculate, a fee that is a fixed percentage of the total budget charged by the implementing institution.

Step two – finding the unit prices

The second step in the process of calculating the budget is to identify the prices for each 'unit' in the budget: salaries by the day or the month, per diems, prices of travel, etc. The amount paid for daily food and lodging has often proved to be a sore point in short-term research projects. If the project hires people who have been given temporary leave of their regular job to participate in the research, they will expect to be paid enough to keep some of the per diem money for themselves. Some expect to keep 50%, others more. In our experience, there is no reason to pay per diems in ways that allow recipients to keep more than half of what is given; nor is there a reason to worry if they keep up half of what they receive. Usually there are many different (and

competing) guidelines not only for which rates to pay, but also the basic principles. In many countries there are official government rates applicable to government employees, another rate used by NGOs, and perhaps a third rate in the private sector. United Nations agencies sometimes have their own rate. Negotiations aiming at finding a compromise between the 'ministerial rates', the 'local university system' and the 'NGO's guidelines' can be very interesting. We strongly recommend that you clarify these issues with the researchers as one of the first things.

Step three – putting numbers down for how many are needed of each unit

The third step involves estimating how much or how many of each of the items in the budget is needed for each activity. At this stage the research team needs to be able to imagine the progress of the project in some detail from beginning to end. The research design will have stated how many days are need for training, for start-up, for contacts, data collection, etc., but it will not specify the need for other things: how much paper is need for questionnaires, how many tape recorders and tapes for interviews, etc.

Step four – summing up

This fourth step involves calculating the totals needed for each category and arriving at a grand total. After this is complete, we recommend that the budget be shown to one or two people experienced in this type of field study to see whether the budget looks plausible and if you have forgotten some item.

General budget considerations

Some may find this procedure self evident. Others may find it overly laborious and detailed. The point we try to make is that by demanding the contractors to go through the steps of this process and making them explicit, one is in a good position to reach at a reasonably sized budget. Thereby, the sponsor is protected from paying over price and the contractor is protected from losing due to under-budgeting.

Program directors will need to examine budgets submitted for their approval. Budgets can be scrutinized to see if they are comprehensive, appropriate, reasonable, and correct. The following questions are usually asked:

- Comprehensive: do the line items in the budget cover all of the project activities? Often budgets do not reflect the amount of time needed for analysis and write-up.

- Appropriate: do the items proposed fit with the activities planned? Is all the equipment requested necessary for the fieldwork and analysis?

- Reasonable: are the amounts suggested reasonable within the local context? Do the salaries to be paid to researchers and interviewers fit within the range people usually receive? Does the per diem adequately cover expenses with some left over? Does the overhead requested seem excessive?

- Correct: are the prices quoted for travel, equipment, and consumable items correct and do the calculations tally?

In conclusion, we want to emphasize the importance of making the budget fit closely with the research design. If the plans for research strategy and specific activities are well presented in the design, the budgeting process will be simple. Occasionally you may find it necessary to modify your activity plans in the light of the budget, either for cost or because something was omitted. An elaborated time schedule for the study is an important reference at this stage. The budgeting process serves as another check on the logic of the design.

In relation to the financial aspects of research, there are certain aspects of a more administrative nature. When and how are the funds to be paid to the contractor? (e.g. expenses for actual costs paid in advance at an biannual basis and the overhead at the completion of the task). How does the auditing take place?

CHAPTER 8

GENERAL CONSIDERATIONS

Ethical considerations

No research project should be planned or implemented without considering its ethical implications. The field is complicated and many ethical codes of conducts have been issued. For biomedical research the most well known are probably *The Declaration of Helsinki* (World Medical Association 2000) and the *International Ethical Guidelines for Biomedical Research involving Human Subjects* (Council for International Organizations of Medical Sciences 2002).

The American Anthropological Association has issued a *Code of Ethics* (1998). The complexity can be boiled down to three questions: 1) are the procedures of *informed consent* satisfactory?; 2) are the project activities in any way *harmful* to the participants?; and 3) who gain the *benefits*?

It should be voluntary for a person to enter or not enter into a study based on adequate information. Acceptance should be based on *informed consent*. It is normal procedure to include an annex in a project proposal which describes the explanation given – often both in English (or equivalent) and local language. The consent may be oral or written according to what is locally appropriate. It seems clear that studies that elicit information from individuals as part of a study are obligated to assure respondents that participation is voluntary, that the results will not become known to others (confidentiality), and that they be informed of any risks and/or benefits for them. In the USA and Europe such assurances take the form of formal declarations and/or texts that respondents read and then sign. In medical studies, Institutional Review Boards (IRB) require that participants be informed of the objective of the procedure, of the risks and benefits, of the possibility of asking questions, and that their participation is voluntary (Singer 1993). The consent may be given verbally or written according to local customs. In societies where few people read it would be better to have an informal conversation than to read a speech (Yoder and Konaté 2002).

Participation in a research project should not entail harm to the participants. In survey contexts as well as in most qualitative research, there is no medical risk involved. However, one should bear in mind, that even some social science data collection methods may be conceived as stressful and thus problematic (e.g. questions probing into traumatic experiences in peoples' lives, which they rather not think about). Another aspect of harmfulness has to do with *confidentiality*. Whatever information (e.g. statements given to the researchers or diagnoses) should be safeguarded and kept inaccessible to other community members or authorities.

Another aspect of research ethics has to do with the derived benefits. Is the respondent getting anything concrete out of the time spent? Will the community benefit at a certain stage? Or have they offered their time so that only others (the researchers or the sponsor or another community) will eventually benefit in one way or another? Do the respondents suffer any loses due to the interview? (e.g. wasted income or reduced harvest). We do not suggest that respondents should be paid for information; such cases should be rare indeed. Whatever compensation is considered, we are convinced that researchers must consider what the local population is giving up (time, labor, reputation, etc.), and that they have a right to a recompense of some sort. We suggest that these issues be considered in other ways.

Example 8.1

> One of us was involved in a project in Eastern Africa in which we followed a school health education program for about two years. Many of the pupils, teachers, and parents were deeply involved and sacrificed quite a lot of their time. For a number of reasons we didn't want to give the individuals any donations. The potential for creating trouble was too big. Some would be envious that they didn't get anything. Some would deserve a bigger share than others. Instead we chose to donate a water tank for roof catchment to each of the two schools. It was a gift to the whole community. Nobody could misappropriate it. Everybody could see it as a token of our gratitude. The tanks were clearly useful to the schools in the study area where water is scarce.

Furthermore, during fieldwork among less privileged groups (which are quite often exactly where the studies are conducted), the researchers will frequently be confronted with cases of people who are in deep trouble but have no resources with which to seek help (illness or hunger). In such cases the researchers should have guidelines (and means) for how to respond (see chapter 7, step 1, 8).

A project proposal should deal with these issues in a serious way - including both the issue of which permits need to be obtained, the balance between benefits and potential harm as well as the proper use of informed consent statements.

Should the study include 'participation'?

A research project may involve the study population in various ways not only as passive sources of information, but as active participants. We will give a brief outline of the general principles without going into details.

During the last few decades, alternative approaches have been developed such as 'Action Research' and 'Participatory Rural Appraisal' (Chambers 1983). In these approaches, the study population takes an active part in the research process. The most common form of participation is that representatives of the community

participate actively in the data collection. Quite often they involve innovative methods such as mapping of villages, transect walks of the local environment, categorisation of certain items (e.g. kinds of food) and subsequent ranking according to preference, which are ideally done by the locals themselves and where the researcher is playing mainly a facilitation role.

In many cases individual from the local population also participate in the data analysis following the data collection. For instance, if a group of villagers are collecting information on infant deaths, the size of farming land, and access to clean water, it may be relatively easy to see patterns of correlation appearing directly from the map drawn in the sand or on paper. If such an approach is applied to evaluate a concrete development project which has been ongoing, it is called participatory evaluation.

The themes of research projects are defined as a rule by agents outside the study population. Typically, they will be defined by the scope of an NGO (e.g. family planning or agro-forestry) or the researchers involved (in case they have been given a grant based on their own independent application). However, a more radical form of participation can be used for needs assessment. Thus, it can be used to enable a community to identify which problems have the highest priority (e.g. schooling for the children, access to clean water or poor health care facilities). However, if such an exercise is initiated, there is of course an ethical dimension in the sense that it is unfair to ask a community to make a broad 'needs assessment', where after they are told that only a certain kind of projects can be funded. In other words, if an NGO has the mandate to work with health and nothing else, that should be stated openly as a limit to the process.

The use of participatory research methods has been contested. The proponents argue that the findings are generally more valid (due to the interest and engagement of the local participants), and that the time span from data collection to the results are available is very short. Furthermore, the engagement of the local population is considered an advantage as is it leads to empowerment and ownership if the study is to be followed up by development initiatives. The critics raise various points. First, it is very difficult to ensure that the "whole" community is involved and has a say in the process. Second, they question whether true participation is at all possible, or whether the researchers are in reality more acting as manipulators rather than facilitators (see example 3.15).

We are not going to make a judgment about these issues. What we find important in this context is that program directors should know that the option exists so that they can request it when appropriate. If they choose to incorporate participatory aspects into a research project, it is important to ensure that the contractor has the necessary expertise with this approach.

Finding out what can go wrong

Research is quite challenging. Unforeseen problems always appear. If researchers try to tell you the opposite, they are wrong. Part of the planning process involves considering what can go wrong, and to imagine potential obstacles and set up means of addressing them. The problems (often called 'critical assumptions' or 'study limitations') fall into several categories.

The first are the ones that we cannot do anything about. Suppose that a civil war breaks out in the area while the study is conducted? Or a natural disaster? Or general elections which are prone to bring civil unrest and preoccupation with politics which may distract local key players? Evidently it is beyond the contractor or the sponsor to do anything about that. But nevertheless it should be considered and written in the proposal. Suppose that we are not talking about potential future events, but real ongoing threatening processes? If that is the case, it may simply not be wise to conduct the study under those circumstances. Or it could be conducted somewhere else? The term 'killer assumption' is used to describe such major factors or events, which may disrupt the project completely.

Another kind of problem is the one that we actually can do something about either as it occurs (e.g. break down of vehicle) or in the planning phase. Especially the latter is important and underlines the need to thoroughly go through the steps outlined in this book.

Example 8.2

> Suppose that we plan to conduct a study which looks into the morbidity of children under five years of age as experienced by mothers in an East African country. We have chosen to make fortnightly visits to a group of households and ask mothers a series of questions in order to get a picture of the reported symptoms of illness. When the project starts, our field assistants, who are the ones administering the questionnaires, come back and seek our advice. They found that in a significant number of households, the mothers are not available. Roughly 25% of children live with one of their grandparents either because their mothers are dead (usually due to the AIDS epidemic), or because in this society children often go and stay with their grandparents. Furthermore, it is the harvest season, so the mothers who do live in the same households as their children are busy in their fields. Consequently, the small children are taken care of by their older siblings (mostly girls). The mothers were simply not there.

The assumption that the mothers were available for interviews proved to be wrong. If we had gained more prior knowledge about the study population, we might have avoided the confusion outlined in example 8.2.

Does the contractor have the sufficient know-how and administrative capacity?

It is difficult to 'buy research'; even more difficult than buying commodities. This is partly because research is complex and partly because you buy the research based on a 'promise'. A research proposal can be seen as a promise to conduct a certain study along with a description of how it will be done. This book focuses on the various phases of the research as described in a proposal. It is often assumed that if consultants know how to write well, they are also able to perform well. This is of course a dangerous assumption. The managers should also assess the contractor's administrative strength and know-how. Consequently the sponsor needs to ask a number of additional questions:

- Do they have the administrative capacity to handle money and manpower of the size in question?
- Do they know the country (region) in which they are going to work?
- Do the key people have the necessary language skills (or are qualified interpreters in place)?
- Do they have the necessary local connections (e.g. ties to local research institutions or NGOs)?
- Do they have competent key persons on board, who can serve as back up for all major activities (e.g. qualitative as well as quantitative data collection methods) – see CVs.
- Is the necessary equipment (e.g. transport, computers or laboratory gadgets) available and does the contractor have the capacity to service them?

The sponsor need to reflect on these basic questions, which are difficult to assess based on the proposal itself. It may be a help (with the consent of the contractor) to talk to sponsors that the contractor has previously worked for.

A bird's-eye view

A research project needs detailed planning. Should the sample size be 100 or 110? Is a questionnaire survey the appropriate method to assess hygienic practices? Is USD 100 an appropriate monthly salary for a research assistant? However, in addition to having an eye for details, a manager must also be able to assess the overall outline of a proposal.

To apply a bird's-eye view. The sponsor should ask questions such as:

- Are the objectives clear and well focussed?
- Are all parts of the general objective reflected in the specific objectives and vice versa? As a rule of thumb the general objective should be equivalent to the sum of the specific objectives.
- Have literature reviews been made on all aspects of the project (i.e. describing state of knowledge regarding all the specific objectives)?
- Do the researchers describe clearly which data collection methods will deal with the various specific objectives, and which sample sizes and procedures apply to them?
- Is the overall design of the project suitable for the objectives it aims to achieve?
- Does it clearly describe the relation of the methods of data analysis to the various data collection techniques?
- Are the project activities all reflected in the time schedule and the budget?
- Is the ethical section taking all parts of the study into consideration?
- Are the critical assumptions adequately described?

A research proposal should appear as a whole. There should be no gaps (e.g. a specific objective that is not covered by the literature review) or overlaps (typically two specific objectives with different phrasing but identical content). Neither should there be anything that falls outside the whole, such as a method that does not refer to one of the specific objectives, or a specific objective that is not encompassed by the general objective. The overall consistency of a project reflects directly on its quality. It matters whether the proposal appears as a consistent whole, or it displays gaps or redundancies. More often than not, serious flaws of project proposals can be seen at this overall level. If a sponsor of research can develop the ability to make such judgments, they are at a great advantage. The importance of the overall perspective cannot be over-emphasized. Experienced reviewers, those who deal with evaluation of project proposals on a daily basis, go straight to the objectives and design and see whether it makes sense. Only if these general issues fit together may it be worthwhile to go into details and read the rest.

ANNEX A

CHECKLIST FOR RAPID CHECK OF RESEARCH PROPOSALS

We suppose that the readers of this book are busy people. Therefore we have made a short list of essential questions that should be borne in mind when evaluating a research proposal. This is the one already shown in the introduction (chapter 1) and which has structured the entire book.

The questions:

Is it an appropriate research question?
Are the objectives clear and well focused?
Are all the key terms clearly defined?
Do we have the information already (literature review)?

The methodology:

Is the study population clearly defined and appropriate?
Is the design clear and easy to follow?
Do the data collection methods fit with the objectives?
Are the procedures for data analysis outlined?
Is the time schedule realistic?

General considerations:

Is the budget sufficiently detailed and relevant?
Are the tasks and the available resources in balance?
Are the critical assumptions considered?
Are plans formulated for the dissemination of the research findings?
Is the study ethically sound?
Is the administrative set-up adequate?
Does the research team have the required expertise?
Applying the bird's-eye view – is the proposal consistent?

ANNEX B

LIST OF DEFINITIONS

Bias: A term for systematic error in study results due to erroneous procedures which distort the findings in a specific direction. Examples: a thermometer, which always measures too high a temperature (measurement bias); a sampling procedure which chooses a sample that favors a special subgroup within a population (selection bias).

Closed: This is an adjective characterizing a question. A closed question is asked in a particular way and determines which answers are allowed. For instance: "have you been to Sudan, yes or no?"

Confidence interval: We can never be 100% sure that a study finding (e.g. a figure) is precisely 'the right one'. There will always be an element of doubt due to random error. A confidence interval is a statistical concept which defines a range within which a given figure is likely to lie.

Confidentiality: A concept related to the ethics of conducting research, whereby the answers of an informant are guaranteed to remain confidential, that is, that no one except the research team will have access to the information provided by the respondent.

Confounder: An underlying factor that can confuse a study that looks at causal relationships. Technically a confounder is an additional cause of the effect being studied and at the same time correlated to the cause being studied. Confounders usually give problems because the sampling procedures did not eliminate them (e.g. by applying random sampling).

Descriptive statistics: A kind of quantitative data analysis that deals with the measurements of the sample (e.g. calculating means) without making any inference about the population from which the sample was taken.

Drop outs: This term is relevant to studies which run over a period of time (e.g. a cohort study). Some of the participants that were originally selected for the study may leave (drop out) for one reason or the other. This is important because it may introduce 'selection bias'.

Emic: From a local, inside point of view; meaning is ascribed to actions or events using local terms and concepts.

Ethno-science: An analytic approach providing an overview of the relevant classification categories of a particular phenomenon as seen from the perspective of the study population.

Etic: From an outside point of view; meaning is ascribed to actions or events using terms and concepts from outside the society concerned. These might be concepts meaningful in another society or concepts derived from social science.

Focus group discussion: A guided group discussion among 6-12 people led by a moderator who has prepared a list of general topics s/he wishes to have the group to discuss. Such discussions are deemed successful if the moderator is able to generate discussion among the group members. In addition an observer should be present to note interactions and non-verbal communication.

Folk taxonomy: Classifications of phenomena as defined by the study population.

Inferential statistics: The kind of quantitative data analysis that allows one to draw conclusions pertaining to the whole population based on the elements sampled and studied.

Informed consent: A term related to research ethics. It refers to the right of people to voluntarily enter or not enter into a study based on adequate information.

Mean: A value used in quantitative data analysis to indicate the sum of all observations divided by the number of observations. It is a synonym of "average."

Median: A value used in quantitative data analysis determined in such a way that 50 % of the observations are smaller than the median and 50% are larger.

Open-ended: A term used to describe questions; it refers to questions that have no precoded or even preconceived answers sought after.

Outlier: A term used mainly in quantitative research referring to a figure that is radically different from other measurements of the same variable, e.g. an extremely low measurement of body weight.

Participation: A strategy for involving a population of a specific area in the activities of a project or programme. This contrasts with a strategy that involves planning all aspects of a programme from the outside.

Power: This term is used in quantitative studies. It describes the study's ability to identify a 'true' difference. The more power the less risk there is to overlook a true difference.

Precision: This term refers to the degree to which any estimate or measurement will be the same when measured over and over.

Proxy: This term describes a variable that stands in the place of another one – usually because it is easier to assess. Examples could be height as a proxy for age or possession of a transistor radio as an indicator of wealth. A proxy may be more or less reliable.

Purposive sampling: This is a sampling procedure where a sample is drawn from a population "with open eyes". The researcher picks the participants based on a set of criteria which should be made explicit. Purposive sampling is mostly used in small qualitative studies. A less suitable term is 'convenient sampling'. See also 'random sampling'.

Random: A synonym of "by chance". A term used primarily to describe selection procedures (sampling) in order to indicate that they are chance driven and not subject to systematic errors (e.g. due to personal preferences of the researcher).

Random sampling: A sampling procedure where the sample is drawn by chance. Random sampling which is mostly used in quantitative studies is a means of eliminating confounders. See also 'purposive sampling'.

Recall bias: A kind of systematic error (bias) that is introduced because data collection, which is based on peoples' memory, is introducing skewed reporting, e.g. they are more likely to remember details in the recent past as compared to distant past.

Representativity: Usually it is not possible to study the total population and a sample has to be drawn. When conclusions are made based on the data collected in the sample, it is always an issue whether they apply to the population from which the sample has been drawn. If that is the case, findings are said to be representative.

Sample: The group of elements (e.g. persons, trees, birthday parties or anything else depending on the research question) that is selected for examination out of a larger group.

Sampling: The procedure by which elements are selected for examination (e.g. interviews or blood sampling) out of the total study population (e.g. the people living in Nebraska).

Sampling frame: The 'list' from within which a selection (sampling) can take place. It can be people holding citizenship in a certain country or trees in a certain forest depending on the research question.

Sampling unit: The element that is subject to the sampling procedure. It can be anything depending on the research question. For instance, if a study is conducted on school education, the sampling unit can be a pupil, a class, a teacher, a lesson, a subject or a school.

Semi-structured: This is a term used to characterize a particular sort of interviews. Such interviews are based on a questioning guide to be used in an individual interview. The guide has an order to the main questions, but allows the respondent to change the order at will.

Selection bias: The systematic error which occurs when the elements that are sampled from the total study population don't all have the same chance of being selected, e.g. that young people are more prone to get enrolled in the study because they are more active.

Significance: This is a figure, a numeric value, which expresses how 'sure' we can be about a quantitative result. Realising that all phenomena are subject to random variation, a certain difference between two values can be real or just appearing because of this random error. The significance value is the probability that a difference occurring in a study is 'not real'.

Statistics: Quantitative data analysis (which can be divided into descriptive and inferential statistics).

Systematic error: See bias.

Triangulation: Triangulation is a term which relates to how a study's methodology can be strengthened thereby making the findings more reliable. It is recommended to: 1) use a variety of supplementary data collection methods; 2) solicit data (e.g. interviews) from various categories of persons; 3) involve different persons for data collection.

Validity: A complicated term with many meanings. Put simple it refers to whether or not a measure of something actually measures what we think it does.

ANNEX C

FURTHER READING

This book is written for beginners and for those who do not normally conduct research themselves, but who nevertheless have an overall responsibility for research projects either as managers or as 'customers'. Consequently we have tried to convey the basic principles with a minimum of technical details. Inevitably this means that we been forced to be superficial at times. In case some of our readers feel like going more in depth, we have recommended some titles below that will give you more skills.

Anthropology

Bernard, H. Russell (2000) *Social Research Methods*. Thousand Oaks, CA: Sage Publications, Inc.

Ember, Carol R. and Ember, Melvin (1999) *Cultural Anthropology*. Upper Saddle River, NJ: Prentice Hall.

Haviland, William (1999) *Cultural Anthropology*. Fort Worth, TX: Harcourt Brace College Publishers.

Epidemiology:

Abramson, J.H. (1988) *Making Sense of Data: A Self-Instruction manual of the Interpretation of Epidemiological Data*. New York: Oxford University Press.

Barker, D.J.P. and Hall, A.J. (1991) *Practical Epidemiology*. Edinburgh: Churchill Livingstone.

Ethics:

American Anthropological Association American Anthropological Association (1998) *Code of Ethics of the American Anthropological Association*. http://www.aaanet.org/committees/ethics/ethcode.htm, date: 10th April 2003.

Bankowski, Z. and Levine, R.J. (1993) *Ethics and Research on Human Subjects: International Guidelines*. Geneva: CIOMS.

Council for International Organizations of Medical Sciences (CIOMS) (2002) *International Ethical Guidelines for Biomedical Research involving Human Subjects.* Geneva: Council for International Organizations of Medical Sciences (CIOMS).

Nuffield Council on Bioethics (2002) *The ethics of research related to healthcare in developing countries*, London: Nuffield Foundation.

Nuffield Council on Bioethics (2004) *The ethics of research related to healthcare in developing countries a follow-up Discussion Paper based on the Workshop held in Cape Town*, South Africa 12-14th February 2004, London: Nuffield Foundation.

World Medical Association (2000) *The Declaration of Helsinki.* 5th revision. Edinburgh: WMA. http://www.wma.net/e/policy/b3.htm. Date: 26th December 2003.

Qualitative data

Briggs, Charles (1986) *Learning How to Ask.* Cambridge: Cambridge University Press.

Dawson, Susan, Manderson, Lenore and Tallo, Verinica (1992) *A Manual for the Use of Focus Groups.* Boston, MA: INFDC.

Kvale, Steinar (1996) *InterViews. An Introduction to Qualitative Research Interviewing*. Thousand Oaks: Sage Publications.

Miles, M. and Huberman, M. (1984) *Qualitative Data Analysis.* Newbury Park, CA: Sage Publications.

Schensul, Jean and LeCompte, Margaret, eds. (1999) *The Ethnographer's Toolkit:* vol 1-7. Walnut Creek, CA: Altamira Press.

Scrimshaw, Nevin S. and Gleason, Gary R. eds. (1992) *RAP: Rapid Assessment Procedures.* Boston, MA: INFDC.

Silverman, David (1993) *Interpreting Qualitative Data.* London: Sage Publications.

Spradley, James P. (1979) *The Ethnographic Interview*. New York: Holt, Rinehart & Winston.

Quantitative data

Campbell Michael J. & Machin, David (1990) *Medical Statistics. A Common Sense Approach*. Chichester: John Wiley and Sons.

McKillup, Steve (2005) *Statistics Explained. An Introductory Guide for Life Scientists.* Cambridge: Cambridge University Press.

Tanur, Judith M. Ed. (1991) *Questions about Questions: Inquiries into the Cognitive Bases of Surveys.* New York: Russell Sage Foundation.

Participatory research and training

Cooke, Bill and Uma Kothari, eds. (2001) *Participation: The New Tyranny?* London: Zed Books.

Koning, Korrie de and Martin, Marion (1996) *Participatory Research in Health: Issues and Experiences.* London: Zed Books.

Mikkelsen, Britha (1995) *Methods for Development Work and Research: A Guide for Practitioners.* Thousand Oaks, CA: sage Publications.

Pathmanathan, Indra and Nik-Safiah, N.I. (1997) *Training of Trainers for Health Systems Research*: Vol. 5. Ottawa: IDRC.

Reporting Techniques

Day, Robert A.(1995) *How to Write and Publish a Scientific Paper.* Cambridge: Cambridge University Press.

Greenhalgh, Trisha (2006) *How to read a paper.* BMJ Books. Oxford: Blackwell Publishing.

ANNEX D

References

Campbell, Michael J. and Machin, David (1990) *Medical Statistics. A Common Sense Approach*. Chichester: John Wiley and Sons.

Cancian, Francesca (1975) *What are Norms? A Study of Beliefs and Action in a Maya Community*. Cambridge: Cambridge University Press.

Chambers, Robert (1983) *Rural Development. Putting the Last First*. Longman Scientific & Technical. Harlow, England.

Dettwyler, Katherine A. (1994) *Dancing Skeletons. Life and Death in West Africa. Prospect Heights*, (Illinois): Waveland Press, Inc.

Deutscher, Irwin, Fred Pestello and Frances Pestello, eds. (1993) *Sentiments and Acts*. New York: Aldine de Gruyter.

Kirk, Jerome and Marc Miller (1986) *Reliability and Validity in Qualitative Research*. Beverly Hills, CA: Sage Publications.

Nabila, J.S. and Clara Fayorsey (1996) *Adolescent fertility and reproductive behavior in Ghana. In* The Changing Family in Ghana, E. Ardayfio-Schandorf, ed. Accra: Ghana Universities Press.

Nyambedha, Erick, Simiyu, Wandibba and Aagaard-Hansen, Jens (2001) *Policy implications of the inadequate support systems for orphans in Western Kenya*. Health Policy, 58:83-96.

Pike, Kenneth L. (1971) *Language in relation to a unified theory of the structure of human behavior*. The Hague: Mouton & Co., 2. revised edition, pp.37-72.

Rifkin, Susan B., Muller, Frits and Bichmann, Wolfgang (1988) *On Measuring Participation*. Social Science and Medicine. Primary Health Care: vol. 26, no.9, pp.931-940.

Singer, Eleanor (1993) *Informed consent and survey response: a summary of the empirical literature*. Journal of Official Statistics 9(2):361-375.

Stone, Linda and Campbell, J. Gabriel (1984) *The use and misuse of surveys in international development: An experiment for Nepal.* Human Organization 43(1):27-37.

UNICEF (1999) *An Inventory of Tools to Support Household and Community Based Programming for Child Survival, Growth and Development.* New York: UNICEF Programme Division.

Yoder, P. Stanley (1995) *Examining ethnomedical diagnoses and treatment choices for diarrheal disorders in Lubumbashi Swahili.* Medical Anthropology 16:211-247.

Yoder, P. Stanley and Hornik, Robert (1996) *Evaluating a radio drama about AIDS in Zambia: assessing program effects.* Studies in Family Planning 27(4):1-16.

Yoder, P. Stanley and Konaté, Mamadou (2002) *Obtaining Informed Consent for HIV Testing: The DHS Experience in Mali.* Calverton, Maryland, USA: ORC Macro.

ANNEX E

LIST OF ABBREVIATIONS

AIDS	Acquired Immune Deficiency Syndrome
ANC	Ante-Natal Care
CHW	Community Health Worker
CI	Confidence Interval
DBL	DBL – Centre for Health Research and Development
FGD	Focus Group Discussion
HIV	Human Immunodeficiency Virus
HMIS	Health Management Information System
IRB	Institutional Review Board
MCQ	Multiple Choice Question
NGO	Non-Governmental Organization
ORT	Oral Re-hydration Therapy
SD	Standard Deviation
TB	Tuberculosis
TBA	Traditional Birth Attendant
UNICEF	The United Nations Children's Fund
VHW	Village Health Worker

ANNEX F

FORMAT FOR RESEARCH PROPOSAL

In this annex we outline the main headings of a research proposal. It should be emphasized that there are different formats and that this is only an example. However, the main points are the same though the terminology varies. The important issue is not what it is called, but that it is described.

Front page (title, investigators, institution, date)

Summary

Introduction (sometimes overlapping with background and literature review, should end with study justification, for qualitative studies theoretical framework)

Objectives (general objective and specific objectives. The general should be equivalent to the sum of the specific)

Methodology

- **Study area and population**
  (where will it take place and among which people?)
- **Data collection methods**
- **Study design**
  (including sampling procedures for *all* data collection methods)
- **Data analysis**
  (for *all* kinds of data collected)

Ethical aspects (informed consent, potential harm, confidentiality, benefits)

Dissemination of results (to all relevant stakeholders)

Time line (describing the various elements of the study from start to end)

Critical assumptions (what can go wrong?)

Budget (including budget justifications)

List of literature

Annexes (for instance question guides, informed consent forms, CVs of researchers)

ISBN 87-91521-06-8